

DOKTORI DISSZERTÁCIÓ

**A GÉPI FORDÍTÁS MINŐSÉGE ÉS  
JAVÍTÁSI LEHETŐSÉGEI**

VARGA ÁGNES

2011

Eötvös Lóránd Tudományegyetem  
Bölcsészettudományi Kar

## DOKTORI DISSZERTÁCIÓ

VARGA ÁGNES

# **A gépi fordítás minősége és javítási lehetőségei**

Nyelvtudományi Doktori Iskola  
A Doktori Iskola vezetője: Dr. Bárdosi Vilmos, CSc habil.  
Fordítástudományi Doktori Program  
A program vezetője: Dr. Klaudy Kinga, DSc habil.

A bizottság tagjai és tudományos fokozatuk:

Elnök: Dr. Klaudy Kinga, DSc habil.  
Belső bíráló: Dr. Heltai Pál, CSc habil.  
Külső bíráló: dr. Kis Balázs, PhD  
Titkár: Dr. Horváth Péter, PhD  
Tag: Dr. Kis Ádám, PhD  
1. póttag: Dr. Pajzs Júlia, PhD  
2. póttag: Dr. Szöllősy Éva, PhD

**Témavezető: Dr. Prószekey Gábor, DSc**

Budapest, 2011

# Tartalomjegyzék

<b>1. Az értekezés témája és célja</b>	<b>7</b>
1.1. Bevezetés	7
1.2. Kutatási célok és tézisek	9
<b>2. A gépi fordításról</b>	<b>11</b>
2.1. Bevezetés	11
2.2. A gépi fordítás típusai	14
2.3. A MetaMorpho fordítórendszer	19
2.4. Mire jó a gépi fordítás?	21
2.5. A fordítástudomány helye a gépi fordításban	27
2.5.1. Vélemények a gépi fordításról	27
2.5.2. Egy egységes elmélet szükségessége	31
2.5.3. A relevanciaelmélet a gépi fordításban	32
2.6. A gépi fordítás helye a fordítástudományban	37
2.7. A tézis eredményeinek összefoglalása	40
<b>3. A gépi fordítás minősége</b>	<b>42</b>
3.1. Bevezetés	42
3.2. $C \times A \times Q$ (meta)elmélet	42
3.3. Teljesen automatikus, általánosan használható fordítórendszerek	43
3.4. A minőség mérésének szempontjai	44
3.5. A minőség mérésének technikái	49
3.5.1. Általános módszerek	49
3.5.2. Az emberi értékelés	50
3.5.3. Az automatikus értékelés módszerei	51
<b>4. Érthetőség és elfogadhatóság</b>	<b>56</b>
4.1. Bevezetés	56
4.2. A géppel fordított szövegek megértése	57
4.2.1. A szövegértés	57
4.2.2. A vizsgálat menete	60
4.2.3. Az érthetőség vizsgálata	62
4.2.4. Az érthetőségi index meghatározása	64
4.2.5. Természetes nyelvű szövegek – emberi fordítások	69
4.2.6. Az eredmények elemzése	70
4.2.7. A gépi fordítás elvárt minősége	71
4.2.8. További szövegek érthetősége	74
4.3. Az eredmények összefoglalása	74
<b>5. Az elfogadhatóság mérése</b>	<b>76</b>
5.1. Bevezetés	76
5.2. Az értékelés módszere	78
5.3. Teszteléshez használt szövegek	79
5.4. Az eredmények kiszámítása	80
5.4.1. A változatok összehasonlítása	83

5.4.2. Az értékelési módszer finomhangolása.....	84
5.4.3. Emberi értékelés – gépi értékelés .....	86
5.5. A METEOR értékelési módszer .....	87
5.6. A METEOR számítási módszer problémája.....	90
5.6.1. A töredezettségi büntetés.....	90
5.6.2. A büntetés alakítása az internetes értékelési módszerhez .....	101
5.7. Eredmények .....	114
5.8. Felmerülő kérdések.....	115
5.9. Az eredmények összefoglalása .....	115
<b>6. A gépi fordítás javításának lehetőségei .....</b>	<b>117</b>
6.1. A gépi fordítás javítása .....	117
6.1.1. A fordítás előtt .....	117
6.1.2. A fordítás alatt .....	118
6.1.3. A fordítás után .....	118
6.1.4. Fordítás vagy utószerkesztés?.....	119
6.1.5. Az utószerkesztett szövegek elfogadhatósága.....	122
6.2. Az automatikus utószerkesztés .....	123
<b>7. Utószerkesztési alapelvek és kategóriák .....</b>	<b>126</b>
7.1. Az utószerkesztés szempontjai .....	126
7.2. A hibák értékelése .....	128
7.2.1. Az utószerkesztés célja.....	128
7.2.2. Hibakategóriák és rendszerek .....	130
7.2.3. Eredmények és elemzés .....	131
7.2.4. Részletes hibakategóriák.....	135
7.2.5. Konklúzió.....	141
7.2.6. Az eredmények összefoglalása.....	142
7.3. Javítási minták .....	142
7.3.1. Az utószerkesztési műveletek automatikus kivonatolása .....	142
7.3.2. Utószerkesztési minták .....	182
7.4. Az eredmények összefoglalása .....	189
<b>8. Az utószerkesztés hatása az érthetőségre és az elfogadhatóságra .....</b>	<b>190</b>
8.1. Bevezetés.....	190
8.2. A minimális utószerkesztés vizsgálata kérdőívvel .....	191
8.2.1. A szövegek érthetősége.....	191
8.2.2. Szubjektív vélemények az utószerkesztett szövegekről.....	193
8.3. A minimális utószerkesztés vizsgálata automatikus módszerrel .....	195
8.4. Az eredmények összefoglalása .....	196
<b>9. Összegzés .....</b>	<b>197</b>
9.1. A tézisek összegzése .....	197
9.2. A kutatás újdonságai, korlátai és kiterjesztési lehetőségei .....	200
<b>10. Irodalomjegyzék.....</b>	<b>203</b>
<b>11. Mellékletek.....</b>	<b>211</b>
11.1. A dolgozatban használt rövidítések .....	211

11.2. A géppel fordított szövegek érthetőségét mérő kérdőív szövegei és kérdései .....	213
11.3. Az abszolút minimálisan utószekesztett szövegek.....	215
11.4. Korpusz .....	217
11.5. 10 szóból álló szegmens összes lehetséges felbontása .....	218
11.6. Az optimális felbontást megtaláló algoritmus C nyelvű kódhoz közeli pseudokódja .....	221

## **Köszönetnyilvánítás**

Nagyon szeretném megköszönni a kitartó biztatást és támogatást a családomnak és a barátaimnak. Köszönöm a hasznos szakmai tanácsokat és ötleteket a konzulensemnek, Prószéky Gábornak és tanárainknak. Nagy hálával tartozom azoknak is, akik segítettek a felmérésekben a kérdőívek kitöltésével és a szövegek javításával, különösen húgomnak, Varga Rozáliának, és barátaimnak, akik rám szánták az idejüket.

# 1. Az értekezés témája és célja

## 1.1. Bevezetés

„A gépi fordítás hasznos? Ez megéri hogy dolgozzon ezen? Ez megéri hogy használja azt?” Sokan, akik még nem találkoztak géppel fordított szöveggel, és hirtelen szembesülnek vele, vagy az emberi fordítás helyett szeretnék használni, gyakran teszik fel ezeket a kérdéseket, azaz hogy érdemes-e a gépi fordítást használni, érdemes-e fejleszteni, és vajon mikor lesz ugyanolyan, mint az emberi fordítás. Ezeket a kérdéseket angolról fordította le a MetaMorpho webfordítója. Látjuk, hogy nem egészen tökéletesek, inkább furcsának érezzük őket, ráadásul nem biztos, hogy elsőre értelmezni tudjuk őket. Hiszen mire is vonatkozik az ez? Ki használjon és mit? Egy kis erőfeszítéssel viszont kitalálhatjuk, mit jelentenek, egy kis változtatással pedig könnyebben érthetőek lesznek: „A gépi fordítás hasznos? Megéri hogy dolgozzunk ezen? Ez megéri hogy használjuk?” Még egy kis javítással pedig már „emberivé” is tehetjük őket: „A gépi fordítás hasznos? Megéri, hogy dolgozzunk rajta? Megéri, hogy használjuk?” A nyers fordítás olvasásához, megértéséhez erőfeszítés kell, nem megy azonnal. Ezt az új „nyelvet” is szinte tanulni kell, és erre nem mindenki hajlandó, tehát a hasznosságával kapcsolatban különbözőek a vélemények. És megéri vajon, hogy fejlesszük, ha egyelőre nem reménykedhetünk abban, hogy valaha is olyan lesz, mint az emberi fordítás? Ahhoz, hogy megkapjuk a választ a kérdésre, el kell képzelnünk, hogy nem tudunk a forrásnyelven, de van egy szövegünk, amelyről gyorsan szeretnénk eldönteni, hogy miről szól, és érdemes-e vele tovább foglalkozni. Például kaptunk egy japán nyelvű e-mailt, de nem akarjuk úgy törölni, hogy nem tudjuk, miről szól. Ekkor van több lehetőségünk: megtanulunk japánul, és ha már a kellő szintre értünk, újra elolvassuk a levelet. Ez rendkívül hosszú és költséges folyamat lenne, és a feladat szempontjából nem célszerű, azonnal el is vethetjük. Másik lehetőségként megkérhetünk vagy megbízhatunk valakit, aki tud japánul, hogy

segítsen. Ha nincs más lehetőségünk, ez megvalósítható megoldás. De ha nincs ilyen ismerősünk, akkor ez a folyamat is időigényes, ráadásul költséges is, ráadásul ezt az időt és pénzt minden ilyen szöveg esetén rá kell fordítani a feladatra. A harmadik lehetőségünk a gépi fordítás, amellyel nem kapunk ugyan tökéletes minőségű szöveget, sőt, lehet, hogy nem is értjük nagy részét a fordításnak, arra azonban megfelel, hogy megtudjuk, miről szól a szöveg. Tehát ebből a szempontból hasznos a gépi fordítás, és megéri fejleszteni, hogy egyre jobb és jobb minőségű, jobban érthető szövegeket kapjunk. Hogyan lehet javítani a minőséget? Mivel mérhetjük a minőség javulását? Egyáltalán mit jelent a gépi fordítás minősége?

Hasznosságuk és sokszor akár érthetőségük ellenére is nagy az ellenérzés a géppel fordított szövegekkel szemben, hiszen annyira eltérnek az általunk elvárttól és a megszokottól. De az idegen nyelven beszélő nyelvhasználata is szokatlan, nyelvtanilag nem helyes, néha érthetetlen és hasonlóságokat mutat a gépi fordítással. Kohéziós eszközök szempontjából pedig bizonyítottan nagyon hasonló a gépi fordításhoz (Mártonyi és Varga 2007: 80-89). Az előzőekben bemutatott kérdéseket egy amerikai anyanyelvű, magyarul tanuló fordította le. Az eredmény a következő lett: „A gép fordítása használos? Megéri dolgozni rajta? Megéri használni őt?” Az ilyen beszélővel szemben azonban legtöbbször hajlandóak vagyunk az erőfeszítésre, és toleránsak vagyunk, ha az üzenet megértése a legfontosabb. Vajon megfelelő célokra használják a gépi fordítást? Vajon változtatható a hozzáállás, és ettől jobban elfogadható lesz a gépi fordítás? Ha a szövegek érthetőek, akkor elfogadhatóak a nyelvtanilag „furcsa”, szokatlan, helytelen szövegek? Ha javítjuk a szövegek minőségét, akkor az elfogadhatóságuk is nagyobb lesz? Mi a kapcsolat az érthetőség és az elfogadhatóság között? A disszertációban ezeket a kérdéseket járjuk körül, ezekre keressük a választ.



## 1.2. Kutatási célok és tézisek

Az értekezés legfőbb célja az angol–magyar, minta alapú gépi fordítás minőségének és javítási, azon belül főként utószerkesztési lehetőségeinek vizsgálata. Szintén célunk, hogy a gépi fordítás és a fordítástudomány között szorosabb kapcsolatot alakítsunk ki. Végül az értekezés – bár objektíven vizsgálja a minőséget, és rávilágít hibáira – a gépi fordítást védelmébe veszi. A következő téziseket bizonyítjuk:

1. tézis. A minőség két fontos szempontja, az elfogadhatóság és az érthetőség a géppel fordított szövegek esetében kapcsolatban áll ugyan, de egymással nem megegyező, és nem felcserélhető fogalom. Azt állítjuk, hogy a géppel fordított szövegek elfogadhatóságának mértéke jóval kisebb, mint ahogy azt az érthetőségük indokolná.

2. tézis. Azt állítjuk, hogy az elfogadhatóság szubjektív kritérium, és számos tényezőtől függ. Ennek ellenére az elfogadhatóságot is mérhetjük objektív módon úgy, hogy ehhez kellően sok olvasói mintát használunk. Az objektív méréshez egy automatikus mérési módszert hoztunk létre.

3. A gépi fordítás első minőségi szempontja az elfogadhatóság. Az elfogadhatóság egyrészt minden más minőségi szempont felett áll. Ha a szöveg nem elfogadható, az egyéb minőségi szempontok sem lesznek relevánsak. Az általunk vizsgált géppel fordított szövegek érthetősége a vártnál magasabb, de ahhoz, hogy a szövegek érthetősége érvényesüljön, és az olvasók profitálhassanak belőle, az elfogadhatóságot is növelni kell. Másrészt az elfogadhatóság olyan szempont, amely az egyéb minőségi szempontoktól függ. Az elfogadhatóságot kétféle módszerrel növelhetjük:

3.1. tézis. A szöveg utószerkesztésével javul a szövegek elfogadhatósága: a hibák kijavításával, azaz a szövegek minimális utószerkesztésével, amellyel az érthetőségük az elfogadhatóságnál még nagyobb mértékben javul.

3.2. tézis. Az elfogadhatóságot az olvasók hozzáállásának javításával is növelhetjük. Azt állítjuk, hogy a relevanciaelmélet a gépi fordítás folyamatát is magyarázza. A relevanciát pedig úgy is növelhetjük, hogy az olvasó hozzáállását megváltoztatjuk, még hozzá úgy, hogy hajlandó legyen nagyobb feldolgozási erőfeszítést tenni a számára szükséges információ kinyeréséhez. Ebben a folyamatban pedig a fordítástudománynak is nagy szerepe van.

4. Az utószerkesztési elvek meghatározásához szükség van a szövegek elemzésére. Az elemzéshez kétféle módszert választottunk, és az utószerkesztésnek két szintjét határoztuk meg.

4.1. tézis. A géppel fordított szövegek kétségtelenül sok hibát tartalmaznak. Ezek a hibák csoportosíthatóak, és súlyosságuk meghatározható az alapján, hogy mennyire gátolják az érthetőséget. A géppel fordított szövegek hibáinak kategorizálásához nem megfelelőek az emberi fordításhoz létrehozott hibakategóriák. Azt állítjuk, hogy a sorba rendezés után legsúlyosabbnak ítélt hibák kijavításával a szövegek érthetősége nagymértékben javul, az emberi fordításét megközelíti. Ezt a javítást hívjuk abszolút minimális utószerkesztésnek.

4.2. tézis. A nagyobb mértékű, minimális utószerkesztéshez szükséges szigorúbb hibakategóriák megállapíthatóak, ha az olvasók javítják ki a szövegeket, majd a javítások alapján mintákat állapítunk meg. Az olvasói javítások automatikus kiszűrése majd kategorizálása alapján létrehozott javítási mintákkal minimális utószerkesztési alapelveket állapítunk meg.

## 2. A gépi fordításról

### 2.1. Bevezetés

Amikor gépi fordításról beszélünk, a kifejezést értelmezhetjük tágabb és szűkebb értelemben. Tágabb értelemben Somers (Somers 1998: 137) alapján a teljesen automatikusan, számítógépes szoftverrel történő fordításon kívül gépi fordításnak nevezhető az interaktív fordítás is, valamint a gépi fordítás magában foglalhatja az elő- és/vagy utószerkesztést is. Ebben a dolgozatban viszont a gépi fordítás terminus kizárólag az elő- vagy utószerkesztés nélküli, teljesen automatikus gépi fordítást jelenti.

Az automatikus gépi fordítás folyamata részben hasonlít, részben pedig nagyon is különbözik a hagyományos, emberi fordítástól. A gépi fordítás ugyanúgy kétnyelvű, közvetítéssel történő kommunikáció, mint az emberi fordítás (Reiß 1978: 27). A kommunikációt szűkebb értelemben vesszük, mégpedig azt a folyamatot, amelyben a forrásnyelvi szövegből megszületik a célnyelvi szöveg. A kommunikáció tágabb értelmezése nem szükséges, hiszen a többi résztvevő nem befolyásolja az eredményt. A rendszer nem tudja figyelembe venni a fordítási szituációt, az olvasót, stb. A fordítási folyamat magyarázatára a gépi fordítás szempontjából elegendő Nida információelméletre támaszkodó elmélete, mely szerint a fordítás kódolási és dekódolási folyamat. A gépi fordításban egyelőre az elsődleges cél a formális ekvivalencia megvalósítása. (Nida 2003: 146, 159). A fordítási folyamatba nem vesszük bele az elő- illetve utószerkesztést.

A gépi fordítás bemenete egy FNY-i szöveg, kimenete pedig egy CNY-i szöveg. Vita tárgyát képezheti, hogy a folyamat kimenete szövegnek minősül-e vagy sem. Szövegnyelvészeti szempontból ugyanis nem tökéletesen elégíti ki a „szövegség” követelményeit. Ezeket a követelményeket Beaugrande adja meg: kohézió, koherencia, szándékosság, elfogadhatóság, informativitás, szituacionalitás, intertextualitás. Beaugrande szerint a ha ezek közül valamelyik

nem érvényesül a szövegre, akkor az illető szöveg nem lesz kommunikatív, tehát ő nem is kezeli őket szöveggént (Beaugrande 1983: 4-10). A géppel fordított szöveg minden szempontnak megfelel valamilyen mértékben, és természetesen fordításoként sem egyformán, de nem tökéletesen. A legcélszerűbb az, ha a kimenetet szövegnek tekintjük, még hozzá úgy, ahogy egy nyelvtanuló által írt vagy fordított levél vagy fogalmazás is szövegnek tekinthető. Ahogy a nemzetközi szakirodalom is, a dolgozatban mi is szöveggént fogunk hivatkozni a gépi fordításokra. Véleményünk szerint a hét szempont közül az elfogadhatóság némileg a többi szempont felett áll, hiszen ha a szöveg az egyéb kritériumok hiányosságainak ellenére elfogadható, akkor az olvasó is szövegnek fogja tekinteni. A kutatásnak pedig pontosan az az egyik célja, hogy megállapítsa, mennyire fogadhatják el az olvasók szövegnek illetve fordításnak.

A gépi fordításban a bemenet és kimenet tehát „ugyanaz”, mint az emberi fordítás folyamatában. Ezenkívül a hagyományos fordítási folyamatban résztvevő szereplők közül az eredeti szöveg szerzője, és az olvasó – akiket az egyszerűség kedvéért nem tekintünk a szűken vett fordítási folyamat részének – ugyanúgy jelen van a gépi fordításban is, de a fordító személye hiányzik. Őt helyettesíti a fordítóprogram. Klaudy szerint a fordítás alkotó tevékenység, hiszen „a fordító minden egyes modat lefordításakor számtalan választás előtt áll, tevékenységének eredménye – a célnyelvi szöveg – számtalan választás, számtalan döntés eredménye. [...] bizonyos objektív törvényszerűségek szerint zajló, ugyanakkor több szubjektív választást megengedő tevékenységről van szó.” (Klaudy 2004: 15)

Newton is azt mondja, hogy a fordítónak szubjektív döntéseket kell hoznia:

“Nevertheless, the human’s ability to infer meaning does not always present translators with clear-cut solutions to translation problems. Producing a translation which may need to convey concepts, attitudes and physical realities alien to the target-language culture demands considerable resourcefulness and creativity. Several acceptable renderings

of a sentence may be proposed but the adoption of one in preference to the others must ultimately be based on subjective criteria.” (Newton 1992: 5)

„Mindazonáltal az ember képessége, hogy kikövetkeztesse a jelentést, nem feltétlenül ad a fordítóknak egyértelmű fordítói megoldásokat. A célnyelvi kultúrától idegen gondolatokat, magatartásformákat és fizikai valóságot átadó fordítás elkészítése jelentős leleményességet és kreativitást kíván. Egy mondatnak számos elfogadható fordítása lehetséges, de ezek közül az egyik kiválasztása végül szubjektív kritériumok alapján kell, hogy történjen.” (Newton 1992: 5)<sup>1</sup>

Láthatjuk tehát, hogy a fordító folyamatosan döntésre kényszerül, sokszor szubjektív döntésre. Hogyan lehetséges tehát még csak gondolnunk is arra, hogy egy gép képes ilyen nehéz folyamat „utánozására”? Hogyan lenne képes a számítógép szubjektív döntéshozatalra, vagy bármilyen alkotó tevékenységre? A gépi fordítás nehézsége pontosan a fordítási tevékenység jellegében rejlik, valamint a folyamat összetettségében.

Mivel innen közelítve a feladat lehetetlennek tűnik, viszont tapasztalatok alapján a gépi fordítás mégiscsak használható, és valóban fordítás – amelynek minősége a különböző rendszerek és különböző nyelvpárok esetében különbözik – a folyamatot célszerűbb más, földhözragadtabb, pragmatikusabb szempontból megközelíteni. Tekintsünk el tehát kezdetben az alkotástól, a döntésektől, és elégedjünk meg azzal, hogy elég, ha egy forrásnyelvi egységnek találunk egy – legalább kezdetben megfelelő, valamilyen szempontból ekvivalensnek tekintett – célnyelvi egységet, még ha az nem is tökéletes, majd próbáljunk meg egyre finomítani a folyamaton. Az eredmény valószínűleg soha nem lesz olyan, mint amelyet egy ember hozna létre, de valljuk be, az emberi fordítások sem tökéletesek, hiszen mi is hibázhatunk, ahogy általában hibázunk is. De nagyon fontos, hogy a gépi fordítást mindig a helyén kezeljük, ne is tekintsük

---

<sup>1</sup> A szó szerinti idézetek fordítása – ha nincs külön jelezve – saját fordítás.

tökéletesnek, ne tekintsük alkotásnak, hanem egy segítő eszköznek, amelyet sok olyan különböző helyzetben használhatunk, ahol a „tökéletes”, de legalábbis sokkal jobb minőségű emberi fordítás nem, vagy csak nehézségek árán áll rendelkezésre.

A fordítási folyamat három részfolyamatból áll: dekódolás, transzkódolás és kódolás (Klaudy 2004: 152). Ezen folyamatok mindegyike külön-külön is rendkívül bonyolult, és ha bármelyikben is hiba történik, az hatással van a fordított szövegre, hiszen a hiba tovább gyűrűzik a következő részfolyamatra. A gépi fordításról is elmondható – a használt algoritmustól és módszertől függően –, hogy nagy általánosságban szintén ugyanebből a három részfolyamatból áll, és a hibák hatása is hasonló.

A következő részben láthatjuk, hogy van olyan gépi fordítási módszer, amely az emberi fordítási folyamatot „utánozza”, és mindhárom részfolyamat fellelhető benne, más módszerek „kihagynak” fázisokat.

## **2.2. A gépi fordítás típusai**

Boitet et al. alapján a gépi fordítást a kutatások különböző korszakaiban különbözőféleképpen csoportosították. Először a '70-es években a gépi fordítási módszereket a fordítás előbb említett három részfolyamatában felhasznált technológia alapján generációkba sorolták (Boitet et al. 2009: 1). Kezdetben, a második világháború utáni időszakban a kutatók – főként a számítástudomány művelői – a mesterséges intelligencia iránti hatalmas lelkesedéssel és naivitással vetették bele magukat a gépi fordításba. A kriptográfia területét felhasználva az első rendszerek az úgynevezett direkt rendszerek voltak, amelyekben a dekódolás és kódolás folyamata helyett egyszerű szótári behelyettesítést használtak. A transzkódolás folyamata kimaradt. Ez a módszer ebben a formában természetesen nem volt sikeres, de a fejlesztésekben és kutatásokban ekkor a nyelvészet szerepét jelentéktelennek tartották, és így nyelvészeket nem is hívtak segítségül. Ekkor az elméletek a formális nyelvekről még nem jelentek meg, a nyelvészek pedig

lehetetlennek tartották, hogy a nyelvet annyira szigorúan és formálisan le tudják írni, hogy a számítógép ebből jó fordítást készíthessen, a számítástudomány kutatói pedig em rendelkeztek megfelelő nyelvészeti ismeretekkel (Somers 2000: 330). Nem csodálkozhatunk ezek után, hogy az 1966-os ALPAC jelentésben a gépi fordítást elmarasztalták és bár kísérletezésre érdemesnek tartották, nem fűztek nagy reményeket a hasznosságához a közeljövőben. Érdemes megnézni, milyen véleményt írtak a gépi fordításról 1966-ban: “Unedited machine output from scientific text is decipherable for the most part, but it is sometimes misleading and sometimes wrong (as is postedited output to a lesser extent), and it makes slow and painful reading” (Pierce 1966: 19). „A tudományos szövegek gépi fordítása nagyrészt megfejthető, de néha félrevezető, néha pedig helytelen (ahogy az utószerkesztett változat is, de kisebb mértékben), és lassú, nehézkes olvasmány” (Pierce 1966: 19). Ha most olvassuk az angol–magyar géppel fordított szövegeket, valószínűleg nagyon hasonló vélemény alakul ki bennünk: a szöveg néha félreérthető, néha helytelen és nehezen olvasható. Ennek ellenére azonban nem szabad arra a következtetésre jutnunk, hogy nem is érdemes vele foglalkozni. Azzal a megállapítással nem értünk egyet, hogy az utószerkesztett változat is ilyen. Meglátásunk szerint az utószerkesztés lényege, hogy a félreértéseket tisztázza, a szöveget pedig érthető formára hozza. A bizottság ítélete a következő volt:

The Committee indeed believes that it is wise to press forward undaunted, in the name of science, but that the motive for doing so cannot sensibly be any foreseeable improvement in practical translation. Perhaps our attitude might be different if there were some pressing need for machine translation, but we find none (Pierce 1966: 24).

A Bizottság hiszi, hogy bölcs dolog a tudomány nevében töretlenül továbbhaladni, de ennek motivációja nem lehet az előrelátható javulás a

gépi fordítás terén. Talán más lenne a véleményünk, ha lenne bármilyen égető szükséglet a gépi fordítás iránt, de ilyet nem látunk (Pierce 1966: 24).

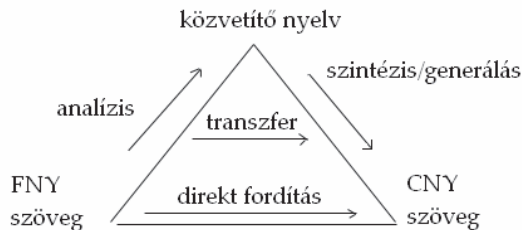
Ezt a véleményt a tapasztalatok már megcáfolták, hiszen például a jól ismert METEO időjárásjelentéseket fordító program már 1977-től működik Kanadában (Thouin 1981: 39). A fordítások rossz minősége ellenére tehát az akkori, gépi fordítást kutatókkal egyetértve úgy gondoljuk, hogy ez a vélemény elhamarkodott és rövidlátó volt (Pierce 1966: 19 , Somers 2000: 330). A gépi fordításnak tehát ez volt az első generációja, a szótári behelyettesítéseken alapuló direkt módszer.

Bár a gépi fordítás kutatása ekkor a finanszírozás hiánya miatt Amerikában és Angliában nagymértékben megtorpant, Európa többi részén azért folytatódott a kutatások. Csak az 1960-as években ajánlotta először Y. Yngve (Boitet 1988: 4), hogy a direkt megközelítés helyett a folyamatot három részfolyamatra kellene bontani: egynyelvű analízis, kétnyelvű transzfer és egynyelvű szintézis, amely megfelel a fordítás három részfolyamatának. Az analízis után kapott reprezentáció egy köztes reprezentáció legyen, és ne a célnyelvre támaszkodjon. Ez a módszer a transzfer megközelítés.

Bar-Hillel 1951-ben ajánlotta, hogy egy közvetítő nyelvet használjanak a köztes reprezentációhoz, például az eszperantót (Bar-Hillel 1951: 235), vagy egy absztrakt közvetítő nyelvet, amelyről azután bármilyen más nyelvre végrehajtható a szintézis folyamata. Schubert is az eszperantót ajánlja erre a célra (Schubert 1992: 79). Ezt a megközelítést hívják interlingua módszernek, és bár a népszerűsége a '80-as, '90-es években volt a legnagyobb, még ma is folynak kísérletek (például Dave et al. 2001: 254). A transzfer és az interlingua módszer a gépi fordítórendszerek második generációja, és ezek gyűjtőnéven indirekt rendszereknek is nevezhetők.

Ezeket a megközelítéseket láthatjuk a következő ábrán, amelyet Vauquois vezetett be 1976-ban a gépi fordítás lépéseinek illusztrálására (Vauquois 1976: 131).





**1. ábra A gépi fordítás első és második generációja**

A '70-es évek generációs felosztása alapján a gépi fordítórendszerek harmadik generációját azok a rendszerek alkotják, amelyek valamilyen formában tudást tartalmaznak a világról, ilyenek a tudásalapú rendszerek (Knowledge-Based Machine Translation). Később majd látni fogjuk, hogy sok kutató feltétlenül szükségesnek tartotta, hogy a gépi fordítórendszer tudással rendelkezzen a világról, és ezt példákkal támasztották alá, de később kiderült, hogy a valóságban, a pragmatikusabb célok eléréséhez erre nincs feltétlenül szükség.

A '80-as években a rendszereket inkább a felhasználók és a felhasználás szempontjából csoportosították: szemlélődők számára, fordítók, szerzők, stb. számára készült rendszerek voltak a különböző kategóriák. De ez a kategorizálás az irodalomban nem terjedt el, míg az első, generációs csoportosítást még ma is használják.

2000-től kezdve a hangsúly átkerült a két fő irányvonalra, a szabályalapú (RBMT, Rule-Based Machine Translation) és a statisztikai megközelítésre (SBMT/SMT, Statistical-Based Machine Translation vagy Statistical Machine Translation). Ezek helyett 1994 óta két másik elnevezést javasol Boitet (Boitet 2009: 2): a szakértői (kézzel készített) és az empirikus (gépi tanulást használó) rendszereket, hogy kifejezze a kettő közti különbséget. Az 1980-as években előtérbe kerültek a korpusz-alapú módszerek a kutatásban, amelyek két fő irányvonala a statisztikai gépi fordítás (SMT) és a példa-alapú gépi fordítás

(EBMT, Example Based Machine Translation). A statisztikai és a példa-alapú rendszerek is kétnyelvű párhuzamos korpuszt használnak.

A statisztikai rendszerek alapvetően a következő modellt használják (Hutchins 2005b: 63): a mondatonként és szavanként is illesztett párhuzamos korpuszból fordítási modellt (a forrásnyelvi és célnyelvi gyakoriságokra) és nyelvi modellt (a célnyelvi szószekvenciák valószínűségei alapján) készítenek, majd minden bemenő szóra a legvalószínűbb célnyelvi szót választják, és meghatározzák ezekre a szavakra a legvalószínűbb sorrendet. A fordítás alapegysége a szó.

A példaalapú rendszerek hasonlóképpen működnek (Hutchins 2005b: 66), de a fordítási egység itt a szó szerkezet (phrase). Az analízis fázisában a bemenő mondatot felbontja megfelelő részekre, ezeket pedig hozzáilleszti a megfelelő forrásnyelvi részekhez az adatbázisban. Ezek a forrásnyelvi részek a minták, amelyek tartalmazhatnak változókat. A szintézis előkészítő fázisában a forrásnyelvi részeket a célnyelvi részekhez illeszti az adatbázisban, valamint származtatja a mintákat/sablonokat (template). A szintézis folyamatában a kinyert célnyelvi részeket átalakítja és összekombinálja, hogy létrehozza a kimenő mondatot. A példa-alapú rendszerek több módszert, technikát is integrálhatnak különböző egyéb megközelítésekéből (szabályalapú, statisztikai, fordítómemóriák.).

A szabályalapú rendszerek nyelvi szabályok, nyelvtanok segítségével elemzik a forrásszöveget, majd ezek alapján generálják célszöveget. A rendszerek természetesen szótárakat is használnak. A fordítás folyamatában a szabályok használatával létrehoznak egy köztes reprezentációt, majd ebből a célnyelvi szöveget.

Boitet et. al. az előzőekben leírt módszerek alapján alapvetően háromféle felosztási rendszert ajánl: a nyelvi felépítés alapján, a technológia alapján és a működési architektúra alapján történő felosztást (Boitet 2009: 2). A nyelvi felépítés szerinti csoportosítás nagyjából megfelel a '70-es évek felosztásának, a technológia

szerinti a 2000-től használt felosztásnak, a működési architektúra szerinti felosztás pedig 4 szempont szerint vizsgálja a rendszereket:

- feladat és felhasználók
- nyelvpárok, mennyiség és típus
- emberi közbeavatkozás szükségességének mértéke
- rendelkezésre álló források

Minden fordítórendszer mind a négy kategóriarendszer szerint kategorizálható, attól függően, hogy melyik szempont az elsődleges az éppen szükséges besoroláshoz.

### **2.3. A MetaMorpho fordítórendszer**

A kutatásban használt MetaMorpho minta-alapú rendszer (pattern-based system) (Prószték és Tihanyi 2002: 19-22), amely az optimumot próbálja megtalálni a példa-alapú és a szabály-alapú rendszerek közötti átmenetben. A rendszer „tudását” minták alkotják, amelyek tulajdonságokkal rendelkeznek. Ha a tulajdonságoknak konkrét értéke van, akkor a mintákat példának nevezzük, ha pedig a tulajdonságok értékei nem „kitöltöttek”, akkor ezek a minták szabályok. Az általánosított minták azok a minták, amelyekben néhány tulajdonság meghatározott. A példákat szótárakból, korpuszokból, kollokációs adatbázisokból generálják, a szabályokat pedig kézzel készítik. Minden forrásnyelvi mintához tartozik egy célnyelvi minta, amelyeket a fordítás folyamatában párosít a rendszer. Előfordulhat, hogy egy célnyelvi egység több mintához is illeszkedik. Ekkor a specifikusabb minták felülírják az általánosabbakat, tehát a rendszer azt a forrásnyelvi mintát választja, amelyben a tulajdonságok értéke jobban kitöltött.

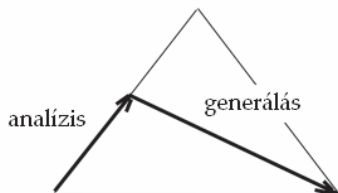
A fordítás folyamatok az analízis fázisa után azonnal szintézis következik a másik nyelven – létrejön a célnyelvi fa –, tehát a transzfer folyamata kimarad. A célnyelvi fából pedig azonnal elkészül a célnyelvi mondat. Az analízis három fázisa:

1. Mondatok tokenizálása szavakra, valamint morfológiai elemzés.

2. A bemenő mondat szintaktikai elemzése alulról felfele működő elemzővel, amely létrehozza a célnyelvi fát terminális és nemterminális szimbólumokkal. Ha a mondat helyes mondat, akkor az elemző egy vagy több gyökérszimbólumot hoz létre.

3. Felépül a célnyelvi fa, majd a fa levelein levő terminális szimbólumokból a morfológiai generátor létrehozza a kimenő mondatot. A terminális szimbólumok halmaza egy véges halmaz, amely azokat az elemeket tartalmazza, amelyekből a nyelv mondatai felépülnek. (Prószéky és Kis 1999: 115)

A fordítás folyamatában nincs szükség transzferre a forrásnyelvi reprezentációból a célnyelvi reprezentációba, valamint köztes nyelvre sincs szükség.



**2. ábra A MetaMorpho fordítási folyamata**

A rendszer unifikációs nyelvtant használ, a tokenizálást követő morfológiai elemzést pedig a HUMOR morfológiai elemző végzi (Prószéky 1994: 213). Elvileg egy kellően bonyolult nyelvtani formalizmussal az adott nyelv szabályai szinte teljes mértékben kifejezhetőek. Akkor miért nem működik a rendszer tökéletesen? A nyelvtani formalizmusokkal szemben három elvárásunk van: nyelvészeti alkalmasság, kifejezőképesség és számítási hatékonyság. (Bővebben lásd Prószéky és Kis 1999: 300-306.) A három követelmény együtt nem valósítható meg maximálisan, mivel kizárják egymást. Így tehát a nyelvi alkalmazások esetében, így a gépi fordítás esetében is, kompromisszumot kell kötni a számítási

hatékonyság javára, és a másik két szempont kárára. Képzeljük csak el, ha minden kivétel, minden egyes kétértelmű lexikai egységet kifejeznénk minták segítségével, milyen hatalmas mintahalmazsal kellene az algoritmusnak boldogulnia (Shieber 1986: 3, fordítás Prószéky és Kis 1999: 300-301). A MetaMorpho a Moose nevű elemzőt és generátort használja, amely szintén ésszerű egyensúlyt teremt a pontosság és a gyorsaság között (Prószéky et al. 2004: 138).

Mivel a nyelvtan esetében tehát kompromisszumot kell kötni, ráadásul a valós helyzetekben a szövegek nem tökéletesek, és nem alkalmazkodnak az ideális helyzetekhez, gyakran előfordul, hogy a mondatot nem sikerül teljes egészében elemezni. Ilyenkor a rendszer összerak egy, a teljes mondatot lefedő mondatrepresentációt a mondat elemzése során megtalált részrepresentációkból. Ezeket a reprezentációkat előzőleg elmenti, és így használja fel. Ezt a fajta fordítást mozaikfordításnak hívjuk (Novák et al. 2008: 113).

## **2.4. Mire jó a gépi fordítás?**

A gépi fordítással kapcsolatos negatív véleményeknek a forrása egyrészt a köztudatban levő, híres „félrefordítások” és példamondatok, amelyek részben legendák (Hutchins 1995: 17-18). Szintén negatív véleményeket kapunk, ha a gépi fordítást olyan célokra szeretnénk használni, amelyekre nem alkalmasak. Ha az interneten elérhető fordítórendszereket kipróbáljuk, meglepően jó és meglepően rossz fordításokat is kaphatunk eredményként. De mitől is függ, hogy a fordítás jó vagy rossz? Mire jó így a gépi fordítás?

Először ki szeretném hangsúlyozni, mire *nem* használható a gépi fordítás. Műfordításra teljességgel alkalmatlan, és ugyanígy alkalmatlan olyan szövegek fordítására, amelyeknél a cél az, hogy kedvező benyomást keltsünk az olvasóban. Ilyen szövegek a reklámok, cégbemutatók, stb. Összefoglalva, ahol maga a szöveg formája a fontos, és nem kizárólag az információ, nem ajánlatos nyers vagy kissé utószerkesztett gépi fordítást használni. Tehát ha Reiß szövegtípusait nézzük,

akkor közülük csak az informatív szövegek gépi fordítása jöhet szóba (Reiß 1981: 163).

A gépi fordítás haszna Sager (Sager 1994: 265) alapján a következő szituációkban a legnagyobb:

- a) Ha nem elég az emberi erőforrás.
- b) Ha nagy az igény a gyors és olcsó fordításra.
- c) Vannak olyan helyzetek és funkciók, amelyekben a gépi fordítás optimálisan teljesít, míg az ember nem. Ezeket a helyzeteket azonosítani kell, és ki kell használni.

Az utolsó pontban említett funkció lehet például a következetes terminológia-használat vagy a helyesírás. Ehhez hozzá kell tennünk még azt a szituációt, amikor a felhasználónak nagyon rövid időn belül (perceken, órákon belül) van szüksége egy adott szövegben található információra. Abban a világban, ahol az internet segítségével másodperceken belül hozzájuthatunk a kívánt információhoz, az emberek egyre inkább hozzászoknak ahhoz, hogy ez az információ minden helyzetben azonnal rendelkezésre álljon. Ilyen helyzetekben jó megoldás lehet a webfordítás. A weben található online fordítórendszerek mindenkinek rendelkezésére állnak – akár ingyen, akár fizetés ellenében -, tehát bárki ellenőrzés nélkül használhatja őket. Emiatt viszont rendkívül fontos a gépi fordítás megbízhatósága és természetesen elfogadhatósága.

A gépi fordítás egymáshoz közeli, hasonló nyelvek esetén is jól használható. Ilyenkor már a nyers fordítás is megbízható minőségű (Kis 2008: 15). A később bemutatott felmérésekből kiderül, hogy a gépi fordítás információszerezésre is használható, egy szöveg témájának azonosítására, valamint annak megállapítására, hogy további, részletes feldolgozására érdemes-e emberi fordítót alkalmazni.

A fordítás minősége és elfogadhatósága a forrásnyelvi szövegtől és a fordítás céljától is függ. Nem mondhatjuk – még a nyers fordításra sem – hogy semmire

nem jó, de azt sem, hogy mindig, minden szituációban jól használható, hiszen sokszor hibás. De „a gépi fordításnak nem kell tökéletesnek lennie ahhoz, hogy használható legyen” (Newton 1992: 4). A nyers gépi fordítást általában nem tekintjük kész terméknek, általában még finomításra, javításra szorul. De a nyers fordítás is használható, ha a cél kifejezetten és csakis az információszerzés (Newton 1992: 4). Ha a kommunikáció a cél- és forrásnyelv különbözősége miatt egyébként lehetetlen lenne, a fordítás pedig nem elérhető, vagy a fordítás túlságosan hosszú időt venne igénybe, ezekben a helyzetekben a nyers gépi fordítás a lehető legjobb eszköz. Például katasztrófa esetében a kommunikáció a nemzetközi mentőcsapatokkal élet és halál kérdése lehet (Bellos 2010: 2).

Jól használható a gépi fordítás egymáshoz hasonló jellegű szakszövegek fordítására is, olyan szövegeknél, amelyeknél az információ a legfontosabb és nem a forma, azaz mint korábban említettük, informatív szövegek fordítására. A gépi fordítás hasznos lehet a fordítási folyamat felgyorsítására is: ha a nyers fordítás egy fordító utószerkeszti, akkor ez az utószerkesztett fordítás helyettesítheti az emberi fordítást. Kérdés az, hogy valóban felgyorsítja-e a folyamatot, valamint hogy ez a feladat könnyebb-e a fordítóknak, és ha még könnyebb is, jobban szeretik-e a fordításnál.

Newton írja, hogy a gépi fordítással közvetlen kapcsolatban levő fordítók hozzáállása pozitív, az elvárásaik realizisztikusak, és ismerik a gépi fordítás korlátait (Newton 1992: 11). Jean Gordon adja meg az ideális, géppel fordítható szöveg jellemzőit:

machine-readable, suitably formatted, limited and stable terminology, without polysemy, of a narrow special subject; no ambiguity, respectful of usage, no typographical or other errors, short sentences, without omission or ellipsis, no complex idioms etc. (idézi Sager 1994: 292)

géppel olvasható, megfelelően formázott, korlátott és állandó terminológiájú, poliszmia nélküli, szűk tárgykörrel rendelkezik,

egyértelmű, a használatot tiszteletben tartja, nincsenek benne szerkesztési vagy egyéb hibák, rövid mondatokkal rendelkezik, nincs benne kihagyás vagy ellipszis, nincsenek benne összetett idiómák stb. (idézi Sager 1994: 292)

Gépi fordítás olvasásakor, használatakor mindig szem előtt kell tartani azt az alapelvet, amelyet Sager mond ki, és elvárásainkat, hozzáállásunkat is ehhez kell alakítani:

To conclude: the basic premise of a theory of MT must be that a machine-translated text is never comparable to a human product of writing or text modification. It has to be considered a product in its own right with its own characteristics. It is the result of a particular automated process chosen deliberately by a writer, an end user, a communication agent or mediator. (Sager 1994: 258)

Összefoglalásképpen: a GF elméletének alapköve az legyen, hogy a géppel fordított szöveg soha nem hasonlítható össze az emberi írással vagy szövegmodosítással. Saját jogán kell terméknek tekinteni saját jellegzetességekkel. Egy sajátos automatizált folyamat eredménye, melyet az író, végső felhasználó, a kommunikációs ügynök vagy közvetítő szándékosan választ. (Sager 1994: 258)

A géppel fordított szöveget mesterséges nyelvű szövegnek kell tekintenünk. A mesterséges nyelv korlátozott, nincs emotív vagy esztétikai jelentése (Sager 1994: 259). Ha a gépi fordítás mégis eredményez ilyen hatást, az nem szándékos, és legtöbb esetben nem az a hatás, amelyet a forrásszöveg kelt. A szövegek olvasásánál ezt fontos tudatosan szem előtt tartani. Ennek a mesterséges nyelvnek saját stílusa van, amely nem ugyanaz, mint a forrásnyelvi szöveg stílusa. A géppel fordított szöveg stílusa komikusnak hat, amelyet főként a szokatlan szerkezeteknek, a furcsa szóösszetételeknek, a nyelvi hibáknak, a szó szerinti fordításnak köszönhet, valamint annak, hogy a szótárból nem a legmegfelelőbb



ekvivalenst választja. Ez a hatás a legszomorúbb, legkomolyabb témájú szövegre is jellemző. A stílus hasonlít a nem anyanyelvre fordító nyelvtanulók stílusához. Nézzünk egy példát.

(1a) Council officers say the Real Madrid star needs permission for barbed wire topping the fence. (ESZ\_001)

(1b) Hivatalosok emberek mondtak, hogy a „Real Madrid“-sztárnak kell kápní egy engedélyet építeni egy tuskést falat. (EF)

(1c) A tanácsi tisztviselők azt mondják, hogy a Real Madrid csillagnak szüksége van arra, hogy engedély a szögesdrótnak felülmúlja a kerítést. (GF\_001)

A hibák természetben van különbség a szövegekben, de a nyelvtanulói fordítások jellege összességében nagyon hasonlít a gépi fordításra. De míg a nyelvtanulókkal szemben türelmesek és toleránsak vagyunk, a gépi fordítással szemben ugyanilyen minőség esetében kritikusak vagyunk, a hozzáállásunk negatív. Ahogy Heltai is állítja (McAlester 1992) és (Campbell 1998) alapján: „A nem anyanyelvre történő fordítás esetében sokszor készek vagyunk elfogadni azt a teljesítményt, amelyet a nem anyanyelvű fordító képes elérni.” De kiemeli azt is, hogy bizonyos típusú szövegeknél nem szabad ilyen kompromisszumokat kötni, például műfordításnál (Heltai 2005: 46). Ha tehát az embertől elfogadjuk a hibákat, a géptől miért nem?

Sager alapján a mesterséges nyelv bizonyos meghatározott célokra megfelelő, funkciója és pragmatikai területe korlátozott. Erről a mesterséges nyelvről feltételezzük, hogy csak olyan elemeket tartalmaz, amelyek a nyelvtanban, korpuszban vagy a szótárban, mintatárban expliciten kifejezettek. A CNY-i elemzés eredménye reprezentálható úgy, hogy teljes mértékben pontosan visszaalakítható lesz a FNY-i eredeti formára. Az ilyen mesterséges nyelv megalkotásánál a fontos és egyben nehézséget jelentő feladatok a következők: a nyelvnek felhasználóbarátnak kell lennie, az információ optimális transzferének meg kell valósulnia, valamint a felhasználó legyen meggyőződve arról, hogy hatékony (Sager 1994: 260). Ezzel Sager tulajdonképpen meg is adja a gépi fordítás minőségének két fontos szempontját: a megbízhatóságot (azaz hűséget) és az

elfogadhatóságot. Ha a CNY-i szöveg a FNY-i szövegben található információt helyesen és optimálisan visszaadja, akkor a GF megbízható. De a megbízhatóság önmagában nem elegendő, ha a szöveget a mesterséges nyelv tulajdonságai miatt a felhasználó nem tartja hatékonynak, azaz számára nem elfogadható. A mesterséges nyelv és a gépi fordítás jellegzetességeit szem előtt tartva sorolja fel Sager a gépi fordításra szánt dokumentumok kiválasztásánál fontos jellemzőket:

- a) Jól fordíthatóak géppel a nagyobb tömegű, több dokumentumból álló, rendszeresen ismétlődő elemeket tartalmazó, homogén szövegek.
  - b) A szöveg legyen géppel könnyen olvasható, megfelelő formátumban. Ehhez hozzátartozik a szöveg szerkesztése, a helyesírás, központozás, szövegformátum stb.
  - c) A szövegben használt terminológia legyen benne a szótárban.
  - d) A szöveg stílusa legyen egyenes, következetes.
  - e) A szövegben ne legyenek gépelési hibák.
  - f) A mondatok, szerkezetek legyenek teljesek, ne legyenek ellipszisek.
- (Sager 1994: 292)

Ezeket támasztja alá Hutchins 5 érve is, amelyekben a gépi fordítás hasznosságát indokolja szinte teljes egészében az előbb felsorolt szempontokra támaszkodva. Tehát a gépi fordításra szükség van, mert először is túlságosan sok a lefordítandó szöveg, másodszor a műszaki szövegek ismétlődő jellegűek és unalmasak, harmadszor a terminológia-fordítás következetességét a gép jobban tudja biztosítani, negyedszer a gépi fordítás növeli a sebességet, ötödször pedig nincs mindig szükség a legjobb minőségű emberi fordításra. A gépi fordítás hasznosságát támasztják alá a gyakorlati példák is, amelyekből látjuk, hogy az olyan szervezeteknél, ahol óriási a fordítanivaló dokumentumok mennyisége – mint például az Európai Unió vagy az Amerikai Légierő –, használják is a gépi fordítórendszereket (Hutchins 2005a: 5).

A valós élethelyzetekben viszont a szövegek nem mindig ilyenek. Még az informatív szövegek sem. Weboldalak fordításakor a szöveg általában rövid, nem homogén, tartalmazhat hibákat, amelyeket a felhasználó legjobb szándéka ellenére sem tud kijavítani, a szövegek bármilyen területről származhatnak, stílusuk különböző, a szerkezetek pedig nem feltétlenül teljesek. Ennek ellenére felhasználóként elvárjuk, hogy elfogadható minőségű szöveget kapjunk. A cél tehát mindig az, hogy a gépi fordítás elfogadható legyen, egyre inkább hasonlítson az emberi fordításra, tehát a minőség mérésekor, a javítást célzó értékeléskor szem előtt kell tartanunk az emberek elvárásait is. A Morphologic Kft. által végzett felmérés alapján a webfordító rendszert leggyakrabban Wikipedia oldalak ([en.wikipedia.org](http://en.wikipedia.org)), majd fanfiction ([www.fanfiction.net](http://www.fanfiction.net)), majd híroldalak ([edition.cnn.com](http://edition.cnn.com)) fordítására használják.

Másik oldalról viszont a felhasználók elvárásait is módosítani kell, fel kell őket világosítani a gépi fordítás lehetőségeiről és korlátairól, arról, hogy mire használható, és mire nem. Véleményünk szerint ebben nagy felelőssége van a fordítástudománynak.

## **2.5. A fordítástudomány helye a gépi fordításban**

### **2.5.1. Vélemények a gépi fordításról**

Mióta a gépi fordítás mindenki által könnyen elérhető lett, sokaknak van róla valamilyen – sokszor megalapozatlan – véleménye. De ebben a részben nem ezekre a véleményekre szeretnénk kitérni, hanem a fordítástudományban megjelenő véleményeket szeretnénk ismertetni.

Egyes megközelítések szerint az elfogadható gépi fordítás nem képzelhető el anélkül, hogy a rendszer valamilyen módon tudással ne rendelkezne a világról, amely a kontextust alkotja. Ezt vallja Hatim és Mason is, akik ígéretesnek látják a szemantikai tudást tartalmazó gépi fordítást. Ők a világról alkotott tudás tárolásában (mesterséges intelligencia) látták a jövőt (Hatim, Mason 1990: 35). Bar-

Hillel – aki egyébként az első teljes munkaidőben a gépi fordítással foglalkozó kutató volt (Somers 2000: 330), – is pont amiatt találta lehetetlennek a gépi fordítást, mert a gép nem rendelkezik a fordító tudásával, és nem ismeri a kontextust.

A human translator, in order to arrive at his high quality output, is often obliged to make intelligent use of extra-linguistic knowledge which sometimes has to be of considerable breadth and depth.

Without this knowledge he would often be in no position to resolve semantical ambiguities. At present no way of constructing machines with such a knowledge is known, nor of writing programs which will ensure intelligent use of this knowledge.

Reasonable goals are then either fully automatic, low quality translation or partly automatic, high quality translation. (Bar-Hillel 1960: 118)

Az emberi fordítónak a jó minőségű kimenet érdekében gyakran fel kell használnia nyelven kívüli tudását, amelynek néha igen mélynek és kiterjedtnek kell lennie.

E nélkül a tudás nélkül gyakran képtelen lenne a szemantikai kétértelműségek feloldására. Jelenleg nem ismerünk semmilyen módszert ilyen tudással rendelkező gépek készítésére, sem pedig olyan programok írására, amelyek biztosítani tudják a tudásnak ezt az intelligens használatát.

Tehát vagy a teljesen automatikus, rossz minőségi vagy a részben automatikus, jó minőségű gépi fordítás elérést tűzhetjük ki ésszerű célként. (Bar-Hillel 1960: 118)

Ez a vélemény elméletben igazolható, mert valóban sok a többértelműség a nyelvhasználatban, és ezek sokszor valóban csak a kontextus segítségével egyértelműsíthetőek. Az átvizsgált szövegek alapján a géppel fordított

szövegekben a legsúlyosabb hibákat a lexikai többértelműség okozza, amely a nyelvtechnológiai alkalmazások egyik fő problémája, de a kutatások fő iránya nem a nyelven kívüli tudás alkalmazása. A MetaMorpho rendszer például többszintű egyértelműsítéssel dolgozik: szófaji, jelentés szerinti és szintaktikai egyértelműsítést végez, valamint a lexikai egységek határainak egyértelműsítését (Gröbler 2004: 89). Bar-Hillel is már felülvizsgálta ezt az említett véleményét, még hozzá pontosan az ellenpéldaként elkészült és működő rendszerek hatására (Boitet 1988: 2).

Már említettük, hogy a kezdeti kutatásokra és a közvélemény korai hozzáállására a mesterséges intelligencia iránti hatalmas lelkesedés és álomszerű, irreális elvárások voltak jellemzőek. Nem is csodálkozhatunk tehát, hogy ezeket az elvárásokat nagy csalódás követte, amely a fent említett ALPAC jelentésben és az angol Lighthill jelentésben (Pierce 1966: 1-124, Russel és Norvig 2000: 65) lett nyilvánvaló. Bár sajnos ezek a jelentések visszavetették a kutatást, mégis jó hatással voltak a gépi fordításra annyiban, hogy a kutatók rákényszerültek, hogy újabb módszereket keressenek, új, egyszerűbb, elérhetőbb célokat tűzzenek ki maguk elé. Az elvárások és a célok tehát később realisztikusabbak, pragmatikusabbak lettek (Boitet 1988: 2). Snell-Horby (Snell-Hornby 2006: 35) is egyfajta pozitívumnak ítéli a rendkívül elmarasztaló, negatív jelentéseket. Ez egyfajta újabb, negatív kezdetet jelentett a gépi fordítás kutatásában, ugyanis pontosan a negatív vélemények miatt került a kutatás és az érdeklődés középpontjába.

Nida is pozitívan nyilatkozik a gépi fordításról. Megemlíti minőségi korlátait, valamint azt is, hogy csak bizonyos típusú szövegek fordítására alkalmas, de ezzel együtt alacsony a költsége, valamint a fordítástudomány számára is hasznos eredményeket hozott a fordítás folyamatáról (Nida 2003: 252). Az ALPAC jelentésben a gépi fordításnak egyedül azt hozzák fel érdekékként, hogy a nyelvészet számára rendkívül hasznos eredményeket hozott.

Early machine translations of simple or selected text, such as those given above, were as deceptively encouraging as "machine translations" of general scientific text have been uniformly discouraging. However, work toward machine translation has produced much valuable linguistic knowledge and insight that we would not otherwise have attained (Pierce 1966: 24).

Az egyszerű vagy válogatott szövegek korai gépi fordításai csalókan biztatóak voltak, ugyanannyira, mint amennyire kiábrándítóak az általános tudományos szövegek 'gépi fordításai'. Mindamelllett a gépi fordítás kutatása nagymértékű nyelvészeti tudást és betekintést adott, amelyet máshogy nem tudtunk volna megszerezni (Pierce 1966: 24).

Heltai a relevanciaelv szemszögéből közelíti a gépi fordítást, azaz arra a kérdésre, hogy „Hasznos-e a gépi fordítás?” a következő választ adja: „A gépi fordítás használhatóságát is a relevancia elve határozza meg: ha nagyon fontos, hogy azonnal kapjunk információt egy idegen nyelvű szöveg tartalmáról, akkor a gépi fordítás is segítséget jelenthet” (Heltai 2005: 46).

Catford 1965-ben a gépi fordítást ranghoz kötött fordításként kategorizálta, mivel a nyelvi szintek között nincs átlépés, a morfémákat morfémákkal, szavakat szavakkal helyettesít, ráadásul magasabb szintű elemekkel a gépi fordító nem boldogul (Catford 1965). Catford véleménye részben még ma is megállja a helyét, és igaz, hogy a gép a ranghoz nem kötött szabad fordításra nem képes, szavaknál magasabb szintű elemeket már tud kezelni, és a különböző szintek között is van váltás. Az idiómák, kollokációk fordítása a rendszer működési elvétől függ, a mintaalapú rendszerben például a minták specifikusságától, statisztikai rendszerben pedig a használt korpusztól.

Nézzük meg a következő példát:

(2a) FNY-i mondat: A gene called p21 **may hold the key to spontaneous healing**, say experts. (ESZ\_046)

(2b) A p21 nevű gén lehet a kulcs a spontán gyógyulás, mondjuk szakértők. (GT)

(2c) Egy gén kiáltott p21 tarthatja a **spontán gyógyítás kulcsát**, mondják a szakértők. (MM, GF\_046)

Az első fordításban kifejezés-kifejezés/tagmondat-tagmondat megfelelés van az egyszerű szó szintű megfelelés helyett, a második fordításban ugyanez a részlet (*tarthatja a ... kulcsát*) szó szinthez kötött fordítás ugyan, de a kiemelt rész itt is átlépi a szó-szó megfelelés szintjét. A mondat szintjét viszont a gépi fordítás nem lépi át. Amit erről Sager 1994-ben leírt, még ma is érvényes: „Since MT is limited to sentence-level transfer, rhetorical and logical structures are not specifically considered by the analysis modules. Any continuity between sentences in target language MT output is therefore purely coincidental” (Sager 1994: 256). „Mivel a GF mondatszintű átvitelre korlátozott, az elemző modul nem veszi kimondottan figyelembe a retorikai és logikai szerkezeteket. Ezért a GF célnyelvi kimenetében a mondatok közötti bármiféle folytonosság csak véletlenszerű.” (Sager 1994: 256)

A statisztikai és a szabályalapú gépi fordításban bizonyos mértékben Catford elméletét láthatjuk, amely szerint a megfelelő számú minta alapján az ekvivalencia-valószínűségek általánosíthatók, és olyan „fordítási szabályok” állíthatók fel, amelyek más szövegekre, illetve „a nyelv egészére” is alkalmazhatók, pontosabban minden olyan szövegre, amely a nyelv egyazon változatát képviseli. Az ekvivalencia-valószínűségeket aszerint állapíthatjuk meg, hogy egy adott FNY-i egységet mekkora valószínűséggel fordítottak ugyanazzal a CNY-i ekvivalenssel (Catford 1965:29-30).

## 2.5.2. Egy egységes elmélet szükségessége

A gépi fordítás interdiszciplináris terület, ráadásul a segítségére levő tudományok annyira távol esnek egymástól, hogy az egyes területekhez közelebb álló kutatók gyakran „elfeledkeznek” a többi tudományágról, amely segítségükre lehetne. Ez a tendencia legélesebben a gépi fordítás kutatásának kezdetén látszott, amikor a fordítórendszereket a nyelvészet bevonása nélkül, csak a technika segítségével

próbálták meg létrehozni. Azóta a helyzet egyre javul és javul, a tudományágak egyre inkább figyelembe veszik mások eredményeit, és új területek is létrejönnek, mint először a természetes nyelvfeldolgozás, majd a nyelvtechnológia. De a fordítástudomány eredményeit és elméleteit még ma is kevesen használják a gépi fordítás kutatásában. Egyetértünk Babych véleményével, miszerint pedig hasznos lenne a fordítástudomány és a gépi fordítás számára is, ha egymás eredményeit használnák, valamint ha közös koncepciókat és elméleteket formálnának (Babych 2005: 8).

Reifler már 1955-ben foglalkozott a gondolattal, hogy egy átfogó fordításelméletet kellene létrehozni a gépi fordításhoz, de ez az átfogó elmélet még nem született meg. Azt is ajánlja, hogy létre kellene hozni egy különálló tudományágot, a gépi fordítási nyelvészetet (MT Linguistics), amely a fordítástudomány egy területe lenne, a gépi fordítás feladataira koncentrálna (Reifler 1955: 21). Ez a tudományág még nem jött létre, de a több területet magába foglaló nyelvtechnológia igen. Szerintünk is rendkívül hasznos lenne, ha a gépi fordításnak szervesebb része lenne a fordítástudomány, mert elméleti és gyakorlati segítséget tudnak nyújtani egymásnak.

### **2.5.3. A relevanciaelmélet a gépi fordításban**

A fordítástudományban Gutt egyetlen, átfogó elméletet és magyarázatot kíván találni a fordításra (Gutt 2000: 24-58). Ez az elmélet a relevanciaelmélet. Vajon ez alkalmazható a gépi fordításra is? Babych azt mondja, hogy a relevanciaelmélet hasznos lehet a gépi fordítás kutatásában is, azaz segíthet a fejlesztésében (Babych 2005: 8). A másik oldalról megközelítve a kérdést az is lehetséges, hogy a gépi fordítás folyamatát megmagyarázza a relevanciaelmélet? Vagy inkább Wilss-szel kell egyetértenünk, aki szerint a fordítási ekvivalenciát nem magyarázhatjuk egy átfogó fordításelmélettel, hanem ezeknek az elméleteknek legalábbis szövegtípusokhoz kell kötődniük, ha nem is egy-egy szöveghez? Kellene a gépi fordításhoz egy külön elmélet?



Véleményünk szerint a relevanciaelmélet megmagyarázhatja a gépi fordítás folyamatát is, ahogyan az emberi fordításét is, még úgy is, ha részleteiben és módszereiben a kettő eltér. A relevanciaelméletben a kommunikációban részt vevő kommunikátor az átadni kívánt üzenet egy részét stimulus formájában közli a közönséggel, aki következtetés útján létrehozza az üzenet egy interpretációját. A következtetés során a közönség szemantikai reprezentációt rendel a nyelvi formához, ezt kiegészíti a kontextus alapján, amelyből kialakul a propozíció forma. A kontextus ilyenkor a közönség kognitív környezete. Ha a kommunikátor és a közönség ugyanazt a kontextust használja, ez a közös kognitív környezetük, amely feltevéseket tartalmaz, amelyekből választanak. A következtetéshez használt erőfeszítés abból ered, hogy a beszélő és a hallgató számára a különböző feltevések nem egyformán könnyen előhívhatók. A hallgatónak pedig a beszélő által szándékozott feltevést kell előhívnia, ami meghatározza az erőfeszítés mértékét.

A hallgató a befektetett erőfeszítéstől bizonyos mennyiségű kontextuális hatást vár, amely lehet az előző feltevéseinek megerősítése, cáfolata, vagy kaphat új kontextuális implikációkat. A beszélő megnyilatkozása akkor lesz a hallgatónak releváns, ha megkapja a számára megfelelő kontextuális hatást a számára elfogadható erőfeszítésért cserébe. A relevancia nem egy abszolút érték, egy skálán kell elképzelnünk. Minél kisebb az erőfeszítés, és minél nagyobb a kontextuális hatás, annál nagyobb a relevancia (Gutt 2000: 24-58).

A fordítás másodlagos kommunikációs szituáció, amelyben a beszélő és a hallgató kontextuális feltevései nem azonosak. Az analitikus feltevések azok a feltevések, amelyek csak a propozíció formából (azaz a gondolat egy reprezentációjából), erednek, a kontextuális feltevések pedig együtt a propozíció formából és a kontextusból erednek (Gutt 2000: 36, Vermes 2004: 7). A kommunikációban megkülönböztetjük a leíró és a interpretív használatot. A leíró használatban a beszélő egy helyzetet ír le, meg van győződve arról, hogy a megnyilatkozásai igazak. Az interpretív használatban a beszélő valaki másnak a

nézeteit mondja el a megnyilatkozásokban anélkül, hogy igaznak vagy hamisnak tartaná őket. Az interpretív használatban az a cél, hogy a megnyilatkozás interpretív mennyire hasonlít egy másik megnyilatkozáshoz. Az interpretív hasonlóság két megnyilatkozás között akkora, amennyire a két megnyilatkozás analitikus és kontextuális implikációi megegyeznek egy adott kontextusban. Az interpretív hasonlóság meglétéhez lényeges, hogy a kontextuális implikációk megegyezzenek, az analitikus implikációk különbözhetnek. Az analitikus implikáció a propozíciós formából következő implikáció, a kontextuális implikáció pedig a kontextusból következő implikáció. A gépi fordításban kizárólag a fordítás interpretív használata képzelhető el (Gutt 2000: 36-37, 61).

A gépi fordítás folyamata teljes egészében leírható a relevanciaelmélet résztvevőivel és folyamataival, azonban a résztvevők tulajdonságaiban, és a folyamatok bizonyos részleteiben vannak eltérések, pontosan a résztvevőkből adódóan. Az egyetlen résztvevő, amely a teljes folyamatot alapvetően megváltoztatja, a fordító helyett dolgozó fordítórendszer. Ahogy a fordító létrehozza a beszélő megnyilatkozásából az interpretációját, vajon a gép is hozhat létre interpretációt? Véleményünk szerint igen, de a gép kognitív környezete az előre megírt szabályokból, mintákból, korpuszokból, nyelvtanokból stb. áll. Az ember intuitívabban, heurisztikusabban hozza létre az interpretációkat, mint a gép, amelynek egyelőre erre nincs, vagy csak nagyon kis mértékben van lehetősége. A gép a saját kognitív környezetéből szintén választ a feltevések közül, és úgy is fogalmazhatunk, hogy számára a feldolgozási erőfeszítés a propozíciós forma feldolgozásából ered elsősorban. A szándékolt és a nem szándékolt feltevések között nem tud különbséget tenni, ez majd a CNY-i olvasó számára fog erőfeszítést jelenteni.

A gépi fordítás különböző elméletei mindig a fordítórendszer működéséhez kapcsolódnak, azaz az újabb és újabb elméletek hatására hoznak létre más és más elven működő rendszereket. Először a direkt rendszereket, később jött a transzfer alapú fordítás, az interlingua elmélet, a statisztikai módszerek, stb. A

relevanciaelmélet összefoglalhatja egy elmélet alá az összes eddigi elméletet. A kommunikáció folyamata és a résztvevők ugyanis minden rendszernél megegyeznek. A különbség magukban a fordítórendszerekben van. Egyrészt a rendszerek kognitív környezetében: ennek jellegében és tartalmában, valamint a kognitív környezet „megtanulása” vagy „beszerzése” is különböző módszerekkel történhet. A másik különbség abból adódik, hogy a rendszer milyen elvek alapján választ a rendelkezésre álló feltevések közül, valamint milyen módszerrel hozza létre az interpretáció(ka)t. A választás és az interpretáció létrehozásának módszere függ a kognitív környezet jellegétől.

A relevancia elmélete a gépi fordításra olyan szempontból különösen jelentős, hogy fordítások jelenlegi minősége miatt az olvasótól elvárt feldolgozási erőfeszítés már akkora lehet, hogy a géppel fordított szöveg a relevancia-skála első határán helyezkedik el, tehát az elfogadhatóság határán mozog. A kérdés az, hogy megéri-e az olvasónak ez a viszonylag nagy erőfeszítés a kontextuális hatás érdekében?

A nehézségek a gépi fordításban egyrészt abból fakadhatnak, hogy a rendszer kognitív környezete nem elég a megfelelő interpretáció kiválasztásához, azaz a rendszernek nincs elég „háttértudása” a választáshoz. Példaként említhetjük, hogy többjelentésű szavak esetén ilyenkor előfordul, hogy a rendszer nem a megfelelő jelentést választja ki (3), vagy nem ismeri fel, hogy köznévről vagy tulajdonnévről (4) van szó.

(3a) Next week it looks like we will get more **showers** and sunshine.  
(ESZ\_057)

(3b) A jövő héten úgy tűnik, hogy több **zuhanyt** és napsütést fogunk kapni.  
(GF\_057)

(4a) Her new boyfriend told her that his father — **Chapman** posed as both father and son — would pick her up in his car and take her to meet him.  
(ESZ\_036)

(4b) Az új barátja elmondta azt neki az apja - a **vándorárus** apaként és fiúként is pózolt - felvinné őt az autójában és elvinné őt, hogy találkozzon vele.  
(GF\_036)

Másrészt szintén hatalmas nehézséget jelent a kognitív környezetek különbözősége. A beszélő és a hallgató kognitív környezete is különbözik, de ez a különbség elhanyagolható ahhoz képest, hogy a gép kognitív környezete mennyiben különbözik mindkettőjük kognitív környezetétől. Ez a különbség hatalmas feldolgozási erőfeszítést követel meg a hallgatótól.

A problémák harmadik oka az lehet, hogy míg az ember az explikatúrákat és az implikatúrákat is értelmezi, a gép csak a propozíciós formából eredő analitikus implikációkat tudja értelmezni, kontextuális implikációk számára egyelőre nem léteznek. Ha a FNY-ben és a CNY-ben a propozíciós formából ugyanazok az implikációk keletkeznek, akkor a hallgató számára csökken a feldolgozási erőfeszítés, ha pedig véletlenül nem, akkor a feldolgozás akár lehetetlenségig nehéz lesz. Ezzel is magyarázható, hogy az egymáshoz jobban hasonlító nyelvek és kultúrák között jobban működik a gépi fordítás (Kis 2008: 15, Callison-Burch et al. 2008: 78).

A szövegek érthetőségének és olvashatóságának nehézségeihez, azaz a géppel fordított szövegek kis mértékű relevanciájához hozzájárul az olvasó kognitív környezete is. Mivel a géppel fordított szöveg részben egy új, mesterséges idegen nyelvnek tekinthető az olvasó szempontjából, feltehető, hogy ha az olvasó ezt az új nyelvet gyakorolja és megtanulja, a szöveg érthetősége és olvashatósága javul. Ezt alátámasztja saját tapasztalatuk, valamint az utószerkesztők visszajelzései is, melyek szerint minél többet olvassuk az egy rendszerrel készült szövegeket, annál jobban megtanuljuk a nyelvet, így egyre kevesebb erőfeszítést kell kifejtünk a megfelelő kontextuális hatás elérése érdekében, tehát a szöveg egyre relevánsabb lesz. A gyakorlással, tanulással az olvasó kognitív környezete egyre több hasonlóságot fog mutatni a „rendszer” kognitív környezetével, tehát a szándékolt feltevéseket az olvasó könnyebben elő tudja hívni.

Bár kontextuális implikatúrákról nem beszélhetünk, amikor a gép fordít, az olvasó számára mégis létezhetnek kontextuális implikációk. Ezek az implikációk magyarázhatják meg azt a jelenséget, amikor a szöveg elemeinek jelentése teljesen

eltér a szándékolt jelentéstől, a kognitív környezet miatt mégis megérti a szöveget, tehát a megfelelő kontextuális hatást kapja meg. Nézzünk meg egy példát:

(5) Azáltal, hogy a kedves melyik időt üli meg az utasülésben, az ajkak annyira szorongtak egy vékony sorozatba nem adni kifejezést annak, hogy a szörnyű vélemények, amiket tartalmaztak a nagy gondolkodásleggömbben, lebegni a feje fölött. (GF\_003)

Ebben a példamondatban az analitikus implikációkat nehéz megtalálni, de a kontextuális implikációk némi erőfeszítéssel kivehetők, így az interpretív hasonlóság az eredeti, FNY-i mondattal valamilyen mértékben megvalósul, a mondat többé-kevésbé érthető lesz, tehát az olvasó kap kontextuális hatást.

Összefoglalva tehát a gépi fordítás folyamatára és általános problémáira megfelelő elmélet a relevanciaelmélet. A részletekben rejlő, a rendszerek különbségeiből fakadó valamint technikai nehézségeket pedig vizsgálni kell külön-külön.

## **2.6. A gépi fordítás helye a fordítástudományban**

Eddig megvizsgáltuk a fordítástudomány helyét és szerepét a gépi fordítás kutatásában, most nézzük meg, hogy a gépi fordítás hol helyezkedik el a fordítástudományban. A gépi fordítással kapcsolatban sajnos rengeteg félreértés és tévedés terjedt el a köztudatban, de még a kutatók körében is, a múlt- és jelenbeli kutatásokkal kapcsolatos általános tudatlanság miatt (Boitet 2009: 1). Hogy vélekednek a fordítástudomány művelői a gépi fordításról? Mi a fordításkutatók feladata és felelőssége a gépi fordítással kapcsolatban?

Szubjektív véleményünk és benyomásaink alapján a fordítástudomány tendenciáiban még mindig a perfekcionista szemléletet követi a gépi fordítással kapcsolatban, és nem vette át a gépi fordítás kutatóitól a pragmatikusabb, praktikusabb, megalkuvó megközelítést. A fordítók természetesen a tökéletesre törekszenek, és a fordítástudományban is ilyen szemmel vizsgálják a gépi fordítást is. Több kutató is megállapította, hogy a fordítók általában hajlamosabbak tiltakozni gépi fordítás használata ellen, mint az egyéb felhasználók (Bowker és

Ehgoetz 2005: 3, Church és Hovy 1993: 249). Ezzel ellentétben viszont az interneten egyre több – angol nyelvű – fordítói bloggal vagy újságcikkkel találkozhatunk, amely a gépi fordítást védelmezi (Bellos 2010, Bell\_Blog 2010). Ezek nem tudományos jellegű művek, bizonyító erejük sem tudományos, de a közvélemény változását mutatják.

Csak hogy a gépi fordítás nem az emberi fordítással akar vetélkedni, és nem is akarja elvenni az emberi fordítók munkáját. Egyszerűen egy segédeszköz lehet olyan helyzetekben, amikor a sebesség döntő fontosságú, amikor az információ kinyerése a legfontosabb, méghozzá gyorsan, amikor hatalmas mennyiségű olyan szövegünk van, amit egyébként nem is fordítanánk le. A közvélemény szintén ugyanezzel a szemlélettel, és nagyítóval vizsgálja a géppel fordított szövegeket, sokszor nem arra a célra használva, amelyre kellene. A negatív vélemény nem teljesen alaptalan, kialakulásához nagyon sok tényező hozzájárul. De véleményünk szerint ez ugyanolyan szűk látókörre és elhamarkodott véleményalkotásra utal, mint az ALPAC jelentés (Pierce 1966: 1-124 ).

Szeretnénk felvázolni a kutatások folyamán informális beszélgetések, benyomások alapján, hogy melyek a negatív vélemények kialakulásához hozzájáruló tényezők:

- Sokan a gépi fordítással először a médiában megjelenő vicces, szélsőséges példákkal találkoznak.
- A weben megtalálható online fordítórendszereket sokan helytelenül használják: egyrészt szótár funkciókat várnak el tőle, miközben a rendszer szövegekre (azaz mondatokra!) van felkészülve.
- A legtöbb figyelmet nem az átlagos, jól lefordított szövegek kapják.
- A felhasználók azonnal tökéletes minőséget várnak, ráadásul az emberek által készített szöveghez hasonlót, sőt, van, aki már szerkesztett, lektorált minőséget vár, amellyel megspórolhatja a hosszadalmas, fárasztó emberi fordítást.

- Sokszor nem akkor találkoznak a géppel fordított szöveggel, amikor szükségük van az eredeti szövegből származó információra.
- Az olvasók sok olyan géppel lefordított szöveggel találkoznak, amelyet nem géppel kellett volna lefordítani, vagy amely legalábbis utószerkesztésre szorult volna. (pl. használati utasítás)

A gépi fordítórendszerek ott vannak az interneten, mindenki által elérhetőek, tehát a fordítástudománnyal foglalkozóknak egyre nagyobb a felelőssége abban, hogy felhívják a figyelmet a gépi fordítás valódi hasznaira, hátrányaira és használhatóságára. Tehát nem hunyhatjuk le a szemünket azzal, hogy a gépi fordítás minősége még nem jó, ne is használjuk.

Felmérések alapján az utószerkesztett gépi fordítást hasznosnak ítélik a felhasználók a kontextus, azaz a felhasználási szituáció figyelembe vételével. Sőt, hasznosnak tartják a még a nem teljesen, hanem csak nagyjából utószerkesztett gépi fordítást is (Bowker és Ehgoetz 2005: 2-4). Ezenkívül mivel az eszköz mindenki számára elérhető, nem lehet megakadályozni a használatát. Ezért sokkal hasznosabbnak ítéljük, ha a fordítástudomány képviselői lennének azok a hiteles források, akik segítenének a megfelelő viselkedéskultúra kialakításában, és a gépi fordítást a megfelelő kontextusba helyeznék.

A feladatok tehát a következők:

- A nagyközönségnek be kell mutatni azokat a feladatokat, amelyekre a gépi fordítás (mindig az aktuális helyzetében és minőségében) megfelelő.
- Ki kell emelni az előforduló hibákat, és a megfelelően kritikus, de nem a teljesen elítélő szemléletet kell hangsúlyozni.
- Figyelmeztetni kell a lehetséges felhasználókat a megfelelő használatra, azaz arra, hogy a gépi fordító nem szótár, és nem is irodalmi fordító.
- Fel kell hívni a figyelmet a jól működő funkciókra, valamint be kell mutatni, milyen jellegű szövegekre használható jól a gépi fordítás.
- A lehető legtöbbet tegyék meg a rendszer jobb működéséért: teszteljék a rendszert, és javasoljanak javításokat.

Ha nem vállaljuk fel ezt a szerepet, akkor több veszéllyel is szembe kell néznünk:

- Egyrészt előfordulhat, hogy széles közönség kezdi használni a nyers fordításokat a jelenlegi formájukban, méghozzá olyan helyzetekben is, ahol egyértelműen nem szabadna. Nem szabad fenntartások és javítások nélkül valódi fordításként használni a nyers fordítást. Ha saját anyanyelvünkre fordítunk, ennek veszélye nem olyan nagy, hiszen látjuk a hibákat. Ha viszont általunk kevésbé vagy egyáltalán nem ismert nyelvre fordítunk, nem tudjuk felmérni a keletkezett szöveg minőségét.
- Ha a gépi fordítást nem megfelelően használjuk, nem vagyunk tisztában a minőségével, használhatóságával, erősségeivel és gyengeségeivel, fennáll annak is a veszélye, hogy elhamarkodottan ítélkezünk, és vagy teljesen jónak vagy teljesen rossznak ítéljük.

Fel kell tennünk azt a kérdést is, hogy vajon milyen hatása lesz a géppel fordított szövegek megjelenésének a nyelvhasználatra és a fordításra?

Véleményünk szerint nagyon könnyen elképzelhető, hogy – ha a nyelvhelyesség szempontjából közelítünk – a hatás nem lesz jó. Előfordulhat, hogy semmilyen változást nem fog előidézni. Az is előfordulhat viszont, hogy változni fog a nyelvhasználat. Ezzel a kérdéssel a jelenlegi tanulmány nem foglalkozik, de mindenképpen megfontolandó kérdésnek érezzük. A változást azonban általában nem lehet kikerülni, de megfelelő igyekezettel – nem pedig ellenállással – jobb irányba lehet terelni.

## **2.7. A tézis eredményeinek összefoglalása**

Annak ellenére, hogy a gépi fordítás nem olyan, mint az emberi fordítás, és nem is helyettesíti, mégis vannak olyan területek, ahol alkalmazható, sőt, hiánypótló szerepe van. Az olvasók elvárásai, a fordított szövegek mesterséges nyelve, a gépi fordítással kapcsolatos tudatlansága miatt az elfogadhatósága nem áll arányban a szövegek minőségével. Megmutattuk, hogy gépi fordítás helyes használatában, az olvasói elvárások realizálásában, a mítoszok és a tévedések eloszlásában a



fordítástudománynak nagy szerepe van, valamint azt, hogy milyen veszélyei lehetnek a helytelen használatnak és előítéleteknek.

Ha a gépi fordítás a fordítástudomány szerves része, akkor szükség van egy olyan elméletre, amely egyaránt alkalmas a gépi és az emberi fordítás folyamatának magyarázatára. Megmutattuk, hogy a relevanciaelmélet lehet ez az elmélet. Annak ellenére, hogy a kétféle fordítási folyamatban a különbség fordító „személyében” rejlik, a relevanciaelmélet megmagyarázza a teljes fordítási folyamatot.

### 3. A gépi fordítás minősége

#### 3.1. Bevezetés

A gépi fordítás értékelése rendkívül népszerű téma, mint azt a szakirodalom mennyisége és szerteágazósága is mutatja. Ebben a dolgozatban nem kívánjuk a teljes szakirodalmat bemutatni, de szeretnénk felvázolni az irányvonalakat, valamint a kutatás szempontjából fontos szempontokat szeretnénk kiemelni. A fordítás javítási lehetőségei szorosan összefüggenek a szövegek minőségével, egyrészt mivel a javulás mértékét valahogyan mérni kell, másrészt pedig az utószerkesztés is a fordítás minőségének mérési eszköze lehet.

Az eddigiekben megállapítottuk, hogy a gépi fordítás bizonyos helyzetekben és bizonyos célokra jól használható. Sőt, bizonyos szempontokból még jobb is, mint az emberi fordítás. A gép nem gépeli el a szavakat, a terminológiai következetessége pedig tökéletesnek mondható (Newton 1992: 5). Már korábban említettük, hogy Sager szerint is a gépi fordítás bizonyos funkciói optimálisak az emberi teljesítménnyel szemben, de ezeket a funkciókat fel kell térképezni (Sager 1994: 265). Fontos tehát, hogy ismerjük ezeket a funkciókat, ismerjük azokat a helyzeteket, amelyekben a gépi fordítás hasznos lehet.

Church és Hovy is azon a véleményen van, hogy a gépi és az emberi fordítás nem egymással versenyző, hanem egymást kiegészítő „eszközök”. Az igazán jól használható gépi alkalmazások elkészítéséhez ki kell használni a számítógép erősségeit, és nem az emberrel kell versenyezni (Church és Hovy 1993: 239).

#### 3.2. C x A x Q (meta)elmélet

Boitet et al. elmélete alapján a gépi fordítás minősége egészen új megvilágításba kerül. A laikus, természetes nyelvfeldolgozásban tapasztalatlan közvélemény azt várja, hogy a GF minden területen jól használható, teljesen automatikus legyen, és a szövegek minősége is tökéletes legyen. De ez az elvárás irreális, ugyanis Boitet et al. C x A x Q (meta)elmélete alapján a három tényező – a lefedettség (coverage),

automatizáltság (automaticity), és a nyelvi minőség (quality) – szorzata mindig jóval kisebb eredményt adhat csak, mint 100%. „Ez a korlátozás a probléma természetéből fakad.” (Boitet et al. 2009: 3) Tehát

$$C \times A \times Q \ll 100\%$$

ahol C a lefedettség, A az automatizáltság mértéke, Q pedig a (nyelvi) minőség. Viszont ha az egyik tényező kárára kompromisszumot kötünk, akkor a másik kettő tényező együtt megközelítheti a 100%-ot. A weben található fordítók például teljesen automatikusak és bármilyen területen használhatók, ennek érdekében viszont a készítőknak a nyelvi minőség kárára áldozatokat kellett hozni. Ez az elmélet segít a kutatóknak és fejlesztőknek döntést hozni a rendszer fejlesztésekor a fordítási szituációnak megfelelően. (Boitet et al. 2009: 3-4)

Ugyanezt a gondolatot fogalmazta meg Church és Hovy már 1993-ban, amikor azt írják, hogy két rendszer „sikerességi mutatói” nem hasonlíthatóak össze, ha az egyik egy szűk területen alkalmazott, kész rendszer, a másik pedig egy még nem teljesen befejezett, általános területeken használható szoftver (Church és Hovy 1993: 242). Ők is, ugyanúgy, ahogy Boitet et al. a feladat-orientált értékelést javasolják. Boitet et. al. alapján az utószerkesztési műveleteken alapuló mérések pontosan ilyen feladat-orientált mérések. (Boitet et al. 2009: 4)

### **3.3. Teljesen automatikus, általánosan használható fordítórendszerek**

A vizsgálatunkban résztvevő fordítórendszer pontosan ilyen webfordító, amely teljesen automatikus, valamint nincsenek területi korlátai. A minőség rovására kötött kompromisszum ellenére – vagy pontosan amiatt –, ahhoz viszont, hogy a fordítás jól használható legyen, a minőségének mégis el kell érnie egy bizonyos mértéket. Ráadásul úgy, hogy nemcsak bizonyos funkciói működnek jól, hanem általánosan jól teljesít, legfőképp pontosság, olvashatóság és érthetőség szempontjából.

Talán már közhelynek is számít, hogy a általában teljesen automatikus gépi fordítás minősége messze elmarad az elvárttól. Ez nemzetközi viszonylatban is

igaz, de már vannak régóta jól és megbízhatóan működő gépi fordítórendszerek, pl. a már említett METEO rendszer (Grimalia és Chandioux 1992: 33-45). Itt viszont a lefedettség nem teljes: a fordított szövegek nyelve korlátozott, mert csak időjárásjelentések szövegeit fordítja, amelyek témájukban és nyelvi szerkezeteikben is korlátozottak.

A magyar nyelvű, interneten elérhető fordítórendszerek fordításait olvasva azt állapíthatjuk meg, hogy a nyers, géppel fordított szövegek minőségén még javítani kell. Egy „kezdő”, gépi fordítást nem ismerő olvasó számára valóban meglepően rossz lehet a szövegek minősége, mivel nagyon eltérnek a megszokott, ember által létrehozott szövegektől. De nagyon fontos, hogy a korábban már említett alapelvet szem előtt tartsuk, és a gépi fordítást saját jogán tekintjük szövegnek, és ne hasonlítsuk össze ember által létrehozott vagy módosított szöveggel (Sager 1994: 258). Hiszen ez a szöveg mesterséges nyelvű szöveg. A gépi fordítástól viszont egészen mást kell elvárunk, mint az emberi fordítástól. Milyen szempontok alapján mérhetjük a minőségét?

### **3.4. A minőség mérésének szempontjai**

Ha a fordítás minőségéről beszélünk, mind a gépi, mind az emberi fordítással kapcsolatban, először fel kell állítanunk a szempontrendszert, amely alapján a minőséget mérjük. Bár a gépi és az emberi fordítás minősége nem hasonlítható össze, bizonyos fő szempontok mindkettőre alkalmazhatóak. De a szempontok és kategóriák megállapításán túl maga a minősítési folyamat, az alkategóriák és az értékelés teljesen eltérő kell, hogy legyen.

Különböző kutatók különböző kategóriákat és minősítési szempontokat írnak le. Nagyon kiterjedt és minden szempontot lefedő értékelési rendszer az Estrella et al. által bemutatott FEMTI, egy kontextuális értékelő keretrendszer kifejezetten gépi fordítórendszerekhez, amelynek segítségével a fordítórendszert a kontextus tükrében lehet értékelni. A kontextus a következőkből tevődik össze:

- mi a feladat,

- ki a szerző,
- milyen típusú a szöveg
- és milyen felhasználóról van szó.

(Estrella 2008: 1)

A keretrendszerben szereplő fő szempontok megegyeznek az ISO/IEC 9126 ISO szabvány által megadott kategóriákkal. A szabványt általánosságban szoftverek minőségének értékelésére hozták létre, így bármilyen szoftver értékelésére használható. A hat fő értékelési kategória:

1. funkcionalitás
2. megbízhatóság
3. használhatóság
4. hatékonyság
5. fenntarthatóság
6. hordozhatóság

Ezek a fő szempontokon belül vannak további szempontok, több szinten. Ebben a dolgozatban minőséget leszűkítjük, először is a fordított szöveg minőségére, majd azon belül is a pontosságra, az érthetőségre, az olvashatóságra és az elfogadhatóságra.

Nem soroljuk fel az FEMTI rendszer összes alkategóriáját, de szeretnénk kiemelni, hogy az általunk vizsgált szempontok közül a pontosság a szabvány alapján funkcionalitás fő kategóriájához tartozik, az érthetőség pedig a használhatósághoz. Mivel a szabvány alapján készült kategóriák nem kifejezetten nyelvi szoftverek minőségére vonatkoznak, és a különböző jellegű rendszereket különbözőképpen kell értékelni, a szabvány alapján az alkategóriákhoz tartozó attribútumokat az egyes rendszertípustól függően szabadon lehet megadni. Az érthetőség itt említett szempontja tehát a szoftver, és nem a szöveg áttekinthetőségére, érthetőségére vonatkozik. Az FEMTI keretrendszerben a kimeneti szöveg érthetősége a *funkcionalitás* kategóriáján belül a CNY-i *megfelelőség* szempontjához tartozik az olvashatósággal, koherenciával és kohézióval együtt. A

nyelvi megfeleléség értékelésére egy külön szempontot hoztak létre szintén a funkcionalitás kategóriáján belül: a *jólformáltság* szempontját. Ide tartozik a morfológia, a központoszás, a lexikai választás és a nyelvtan. A *pontosság*hoz a terminológia, hűség és a következetesség tartozik (Hovy et al. 2002a: 1239-1241, Hovy et al. 2002b: 43-59, FEMTI 2002).

Ebben a szempontrendszerben az elfogadhatóság egyáltalán nem szerepel. Úgy látjuk, hogy eszerint az elfogadhatóság egy külső, minden szempont felett álló kategória, amelyet megközelíthetünk felülről és alulról is. Felülről nézve az elfogadhatóság egy legelső, mindennek feletti szempont, amely nélkül a rendszer értékelésének nincs is értelme. Ha a rendszer (és az általa létrehozott szöveg) nem elfogadható, felesleges is értékelni. Alulról közelítve pedig az elfogadhatóság az értékelés végső eredménye lehet, amely a sok szempontra adott értékelésből tevődik össze.

Fiederer és O'Brien Hutchins és Somers szempontjait használják értékelési vizsgálatukban: egy szempont a hűség vagy pontosság, a következő az érthetőség vagy világosság, a harmadik pedig a stílus (Fiederer és O'Brien 2009: 54-55). Mindhárom szempont csak a szövegre vonatkozik. Arnold a felhasználó, azaz fogyasztó szempontjából írja le a minősítési szempontokat, tehát olyan szempontokat ad meg, amelyek segítenek dönteni abban, hogy a fogyasztó a piacon megtalálható termékek közül melyiket válassza. Ezek a szempontok:

- Változások támogatása – ha a munkatársak nem támogatják a gépi fordítórendszer bevezetését, akkor ennek rájuk kényszerítése egészen biztosan rossz eredményt hoz.
- Technikai környezet – az új rendszer által használt formátumok legyenek összhangban a többi rendszer (szövegszerkesztő, stb.) által használt formátumokkal.
- Az eladó cég stabilitása – a későbbi támogatás érdekében jó, ha az eladó cég megfelelő piaci helyzetben van.

Ezen külső szempontok után következik a rendszer teljesítményének vizsgálata: a sebesség és minőség alapján. A minőségben belül pedig vizsgálni kell a korábban említett szempontokat: érthetőség és pontosság vagy hűség (Arnold 1994: 158-160). A minőség mérésekor a relevancia elvét is szem előtt akarjuk tartani, még hozzá úgy, hogy a szövegnek nem közvetlenül a helyességét vizsgáljuk, hanem elegendőnek tartjuk, ha megtaláljuk azt az egyensúlyt, ahol a feldolgozási erőfeszítés és a kontextuális hatás összhangban van. Ezért mérjük az elfogadhatóságot és az érthetőséget, és ezért célozzuk a minimális utószerkesztést.

Kutatásunk szempontjából a minőség mérésének külső szempontjai nem tűnnek relevánsnak, mivel a dolgozatban a minőség több összetevőjéből csak a szövegre koncentrálnunk. De a teljességhez hozzátartoznak ezek a szempontok is, és fontosnak tartjuk ezeket is kiemelni.

A gépi fordítás előnyei miatt (sebesség, alacsony költség) és mivel hatalmas az igény a fordításra, érdemes a gépi fordítás minőségével és javításával foglalkozni. De Sager érvei miatt az emberi tényezőt olyan szempontból szabad csak figyelembe venni, hogy a géppel fordított szöveget is emberek fogják olvasni. Így tehát az emberi fordításokhoz létrehozott értékelési módszerek helyett – mint például a Dróth által leírt szempontrendszer (Dróth 2002: 98) – a gépi fordításhoz saját szempontrendszerek készültek.

A szakirodalomban szereplő szempontrendszereket összefoglalva az értékelési szempontok között feltétlenül szerepelnek a külső minőségi tényezők: például a rendszer rugalmassága, költsége, funkcionalitása stb., a kontextus: a szerző, feladat és a szövegtípus, valamint magára a szövegre vonatkozó tényezők (Estrella 2008:1). A szövegre vonatkozó szempontok a különböző szerzőknél kissé eltérnek, de mindegyiküknél szerepel a pontosság, az érthetőség és az elfogadhatóság (Fiederer és O'Brien 2009: 54-56, Arnold 1994:158-160, Hutchins et al. 1992: 163). A pontosság a FNY-i és CNY-i szöveg közötti megfelelést jelenti, az elfogadhatóság pedig azt, hogy az olvasó el tudja-e fogadni az adott fordítást szöveggént. Az érthetőség magában foglalja a szöveg folyékonyágát, nyelvtani,

lexikai pontosságát. Boitet hangsúlyozza, hogy a gépnek nem kell az ember viselkedését utánoznia, az egyetlen fontos dolog, hogy az emberek számára elfogadható fordítást adjon (Boitet 1988: 2). Sager minőségi kategóriái: hűség, érthetőség és hasznosság (Sager 1994: 197). A nagyobb kampányokban, amelyekben a géppel fordított szövegek minőségét mérték, a két fő szempont a pontosság (adequacy) és az gördülékenység/olvashatóság (fluency) volt (Callison-Burch et al. 2008:72, Koehn és Monz 2006: 105). Gamon et al. szintén ezt a két szempontot használja, de kiemeli, hogy az olvashatóságot könnyebb mérni (Gamon et al. 2005: 104).

Azt már korábban leszögeztük, és alapul vettük, hogy gépi fordítás kimenete, a fordítás szöveg. Azt nem állítjuk ezzel, hogy teljes mértékben kommunikatív szöveg, mivel a szövegek jellemzői közül általában egyik sem érvényesül maradéktalanul a nyers gépi fordításban. Azt szeretnénk tehát megvizsgálni, hogy ezek a szövegek mennyire felelnek meg a Beaugrande által megadott hét jellemző közül az elfogadhatóság kritériumának (Beaugrande 1983: 4-10). Kutatásunkban azért az elfogadhatóságra koncentrálunk, mert véleményünk szerint az elfogadhatóság részben kompenzál a hiányosságokért a másik hat kategóriában. Mivel a gépi fordítás bizonyos szituációkban, szándékosan választott segédeszköz, így nem feltétlenül várjuk el, hogy a szöveg jellemzői közül mindegyik teljes mértékben érvényes legyen rá. Tehát egyetértünk Boitet fent említett megállapításával, és azt az álláspontot képviseljük, hogy az elfogadhatóság és a megbízhatóság az első két legfontosabb szempont a minőség mérésekor. Természetesen ez nem tökéletességet jelent, és összhangban van a gépi fordítás mai kutatási irányzataival, melyek az álomszerű tökéletesség helyett realisztikusabb részcélokat tűz ki maga elé, majd ezek elérésével újabb részcélok következnek a minőség javítására.



### 3.5. A minőség mérésének technikái

#### 3.5.1. Általános módszerek

A gépi fordítás minőségének mérésére – jellegéből adódóan – az emberi fordítás értékelésétől teljesen eltérő és a hagyományos értékeléshez képest szokatlan módszerek is léteznek. A mérés jellegét és módszerét természetesen megszabják a mérési kritériumok és a mérés célja. Mostantól kezdve a dolgozatban csak a szöveg minőségével foglalkozunk, és nem térünk ki az egyéb – fent említett – külső szempontokra. Ha a minőségre hivatkozunk, mindig csak a szöveg minőségéről beszélünk. A szöveg minőségét alapvetően kétféleképpen mérhetjük: automatikusan vagy emberi értékeléssel. Az automatikus módszereket többé-kevésbé objektívnak nevezhetjük, az emberi értékelés pedig mindig valamilyen mértékben szubjektív.

Az emberi értékelés kétféleképpen történhet: sorrend felállításával vagy a kijelölt szövegrészek értékelésével. Az automatikus értékelés vagy valamilyen ember által létrehozott referenciaszöveghez hasonlítja a géppel fordított szöveget, vagy pedig egy létrehozott formális szempontrendszer alapján, referenciaszövegek nélkül értékeli a szöveget, de ezek ötvözésére is van példa. A mérés jellegétől vagy technikájától függetlenül – ki és hogyan végzi az értékelést – Chuch és Hovy három alapvető módszert ír le a szöveg-alapú értékelésre, amelyek 1993-ig jellemzőek voltak:

1. Mondat-alapú mérés

Valamilyen szempont szerint vizsgálja a helyes mondatok arányát.

2. Érthetőség-alapú mérés

A szöveg érthetőségét, görbülékenységet vizsgálja.

3. Utószerkesztési műveletek mennyisége alapján történő mérés (Church és Hovy 1993: 4)

Azt méri, hogy mennyi utószerkesztési művelet szükséges ahhoz, hogy a géppel fordított nyers szöveget valamilyen funkcióra alkalmas szöveggé alakítsuk.

A jelenlegi mérési módszerek többé-kevésbé besorolhatóak ebbe a három kategóriába, de a szakirodalom alapján általában nem ez a felosztás jellemző.

Boitet et al. alapján inkább a referenciafordítást használó, a szubjektív és a feladatorientált értékelési felosztás a jellemző (Boitet 2009: 4), annak ellenére, hogy léteznek olyan módszerek, amelyek nem sorolhatóak be ebbe a három kategóriába. Vannak például az olyan, automatikus, mondat-alapú értékelési módszerek, amelyek nem használnak semmilyen referenciaszöveget, pl. (Gamon et al. 2005: 103-111). Ezek a módszerek besorolhatóak Church és Hovy 1-es kategóriájába, annak ellenére, hogy esetleg nem a helyes mondatok *száma* a lényeges.

A gépi fordítás minőségének mérése lehet külső vagy belső jellegű. A belső mérés a rendszer technológiájának figyelembe vételével történik, általában fejlesztési célokból. Ilyen mérési módszerek és eredménye kevesebbet szerepelnek kutatásokban. A külső mérési módszerek a rendszert fekete doboznak tekintik, és nem foglalkoznak a használt technológiával (Boitet 2009: 4). A dolgozatban szereplő kutatásokban is külső mérési módszereket használunk, de mindig ugyanazt a fordítórendszert.

### **3.5.2. Az emberi értékelés**

A minőség mérésére a legkézenfekvőbb és bár a legköltségesebb, leglassabb és legszubjektívebb, mégis a legmegbízhatóbb módszer az emberi „kézi” értékelés. Ezek a szubjektív értékelési technikák a rendszer jövőbeli használhatóságának megbecslésére használhatóak (Boitet 2009: 4). Ez a hagyományos fordításban is megszokott értékelési módszer, azaz az értékelő személy elolvassa és értékeli, esetleg kijavítja a szöveget adott szempontok alapján. A módszer sajnos szubjektív, mivel általában egy ember végzi az értékelést. Sager szerint legalább három semleges fél szükséges a megbízható értékeléshez (Sager 1994: 198). Hamon et. al. szintén ajánl módszereket a megbízhatóság növelésére vagy biztosítására, az emberi értékelés ugyanis valóban következtelen és megbízhatatlan. Mindenképpen több értékelőt kell alkalmazni, de az értékelés után még további finomításra is szükség van, azaz mérni kell az értékelők

megbízhatóságát, valamint törölni a nem megbízható értékelők értékeléseit (Hamon 2008: 21). A megbízhatóság biztosítására minden vizsgálatban több értékelő értékelését használjuk.

Az általunk vizsgált értékelési szempontok közül (pontosság, érthetőség, olvashatóság, elfogadhatóság) az elfogadhatóság a legszubjektívebb szempont, hiszen az értékelő sokszor nem racionálisan, hanem érzelmei, aktuális lelki állapota alapján dönt, hiszen az elfogadhatóságra nehéz pontosan meghatározható, objektív értékelési kategóriákat létrehozni. Mégis szükség van ezekre a szubjektív vizsgálatokra, hiszen a felhasználók az életben is így döntenek. A méréskor viszont a fent említett szempontok miatt biztosítani kell a megbízhatóságot.

Doyon et. al. olyan felmérést készített, amelyben az utószerkesztés hatásának vizsgálata mellett felmérték a géppel fordított szövegek elfogadhatóságát. A fordítók és elemzők által adott válaszokban a nyers gépi fordítás minden esetben az elfogadhatósági szint alatt maradt (Doyon 2008: 1-8). Bowker és Ehgoetz. hasonló vizsgálatot folytattak egyetemi oktatók körében. Ők közvetetten vizsgálták az elfogadhatóságot azzal, hogy az emberi fordítás, nyers gépi fordítás és az utószerkesztett gépi fordítás közül kellett a legmegfelelőbbet választani. A választáshoz három tényezőt kellett figyelembe venni: a minőséget, költséget és a sebességet. A felmérés eredménye az lett, hogy a gépi fordítást egyetlen alany sem találta elfogadhatónak, mivel hibákat tartalmazott, és „az olvasóknak túl sok időt vett igénybe a szándékolt jelentés kitalálása” (Bowker és Ehgoetz 2005:4, fordítás tőlem). Ez a megállapítás teljesen összhangban van a relevancia elméletével. Az alanyok azt jelzik itt, hogy a feldolgozási erőfeszítés túl nagy a kontextuális hatásért cserébe.

### **3.5.3. Az automatikus értékelés módszerei**

A gépi fordítások értékelésére a másik, kevésbé költséges, objektívnek nevezhető módszer a gépi értékelés. Mivel géppel létrehozott szövegekről van szó, ahol a

sebesség fontos tényező, a szövegek értékelésénél is hasznos, ha ez a folyamat minél gyorsabb. Az automatikus mérési módszerek haszna, hogy „olcsó, gyors, megismételhető és objektív értékelést” biztosítanak (Popescu-Belis 2003: 309). 2001-2002-ben kezdtek elterjedni ezek az – általában – referenciafordításon alapuló objektív külső mérési módszerek (Boitet 2009: 4). Objektívnek nevezhető a gépi értékelés annyiban, hogy a kiválasztott szempont alapján a szövegeket elfogultság nélkül, pillanatnyi körülményektől és személyes preferenciáktól függetlenül értékeli. Az értékelő módszer és az értékelendő szempontok kiválasztása azonban a módszer létrehozójától függ, tehát ennyiben ez a módszer is részben szubjektív. Gamon et. al. alapján az értékelés két célt szolgál: a különböző rendszerek összehasonlítását, és a hibák feltárását a javítás, fejlesztés érdekében (Gamon et. al. 2005: 104). A gépi értékelés rendkívül hasznos, mert az ember kis minőségi változásokat nem vesz észre, a gépi értékelő módszer pedig egészen apró javulást is képes jelezni. Ezt alátámasztja Boitet et. al. is, amikor azt mondja, hogy a referenciafordításokat használó módszerek a gépi fordítórendszer javulásának mérésére alkalmasak (Boitet 2009: 4).

Az automatikus módszerek egy számot rendelnek a fordításhoz, amely a szöveg minőségét jelenti.. Banjeree et al. (Banjeree et al. 2005: 67) öt szempontot sorol fel, amelynek egy jó értékelés módszernek meg kell felelnie: összhang az emberi értékeléssel (correlation), érzékenység (sensitivity), következetesség (consistency), megbízhatóság (reliability), általánosság (generality). Ezek közül azonban végső soron az értékelési metrika összhangja a megbízható emberi értékeléssel lehet a koherencia szinte egyetlen bizonyítéka (Popescu-Belis 2003: 8)

Rengeteg automatikus értékelési módszer létezik, ezek közül néhányat szeretnék csak bemutatni. A kizárólag statisztikai alapon működő metrikák általában emberi referencia fordítást használnak az értékeléshez, de vannak kísérletek a referencia fordítások kiküszöbölésére. Talán a legismertebb és legszélesebb körben használt módszer a BLEU. Ez a módszer az értékelendő szöveg három emberi referencia fordítását használja. Az értékelésre kijelölt

szöveget n-gramokra bontja, azaz szavakra, szópárookra, szóhármásokra és – négyesekre. Minden n-gramot ezután megkeres a referencia fordításokban, majd a találatok alapján kap a szöveg egy 0 és 1 közé eső indexet, amely a minőségét jelzi. Minél nagyobb ez a szám, az értékelendő szöveg annál jobban hasonlít az emberi fordításokhoz, tehát annál jobb a minősége (Papineni et al. 2002: 311-318). Fontos megjegyezni, hogy ha emberi fordítást értékelünk ezzel a módszerrel, az sem fogja elérni az 1-es értéket, hiszen az emberek sem hoznak létre azonos fordításokat.

Az NIST módszere nagyon hasonló a Bleu módszerhez, de itt a találatokat súlyozzák. Minél ritkábban fordul elő egy n-gram, annál nagyobb súlyt kap, mert ennek nagyobb a jelentősége (Doddington et al. 2002: 128-132). Coughlin egy mérési teszt alapján nem talált erős indokot arra, hogy a kettő közül bármelyiket előnyben részesítse, annyira hasonló eredményeket hoztak (Coughlin 2003: 65). A METEOR metrikát a Bleu módszer gyengeségeinek megoldására hozta létre Banjeree és Lavie. A Bleu a fedés (recall) értéket csak közvetetten veszi figyelembe az index kiszámításakor, a METEOR pedig 9-szeres súllyal számítja bele a harmonikus közép kiszámításakor. A Bleu a nyelvi jól-formáltság mérésére hosszabb n-gramokat használ, míg a METEOR csak unigrammokat illeszt. Valamint a Bleu nem a igényel teljes egyezést a szavak között, ami helytelen illesztéseket eredményez gyakran előforduló szavak esetén, és geometriai átlagot használ, amely 0 értéket ad, ha az n-gramnak akár csak egy eleme nem illeszkedik. A METEOR szinonimákat és pontos találatokat is illeszt, valamint megkeresi az egyező szótöveket is, valamint bünteti a nem megfelelő szórendet illetve a töredezettséget (Banjeree és Lavie 2005 :67-70, Agarwal és Lavie 2008: 115-118).

A TER (Translation Edit Rate, fordítási szerkesztési műveletek aránya), vagy HTER (Human Translation Edit Rate) a szerkesztési távolságot méri az értékelendő fordítás és a referencia fordítás között. A szerkesztési távolság azoknak a szerkesztési műveleteknek a száma, amelyeket el kell végezni ahhoz, hogy az értékelendő szöveg megegyezzen a referencia szöveggel. A szerkesztési műveletek: beszúrás, törlés, áthelyezés, átirás; ezek mindegyike egy műveletnek

számít. A TER értéke a szerkesztési műveletek száma osztva a referenciaszavak átlagos számával. A TER tulajdonképpen megegyezik a WER módszer elvével (Word Error Rate, hibás szavak aránya), amely a beszédfelismerési kutatásokból származik, valamint a Translation Error Rate módszerrel, de Snover et al. megfelelőbbnek tartja a Translation Edit Rate elnevezést. Az eredmények alapján a (H)TER jobban korrelál az emberi ítélettel, mint a BLEU (akár annotált referenciával) vagy a HMETEOR (Snover 2006: 223-231).

Coughlin 2003-as vizsgálata alapján a BLEU index nagyban, pozitívan és lineárisan korrelál az emberi ítélettel a vizsgált adathalmazra és fordítórendszerekre vonatkoztatva (Coughlin 2003:69). Később viszont kiderült, hogy bár a statisztikai gépi fordítás erősen támaszkodik a BLEU módszerre az értékelésben, Callison-Burch et. al. alapján értékeléskor óvatosságnak kell lenni a következtetések levonásában. A kutatás alapján ugyanis a gyakorlatban a jobb BLEU érték nem feltétlenül jelent jobb minőségű fordítást is. Szintén megtevesztő lehet, ha különböző elveken működő fordítórendszereket hasonlítunk össze BLEU index alapján, ugyanis az eredmények elemzése azt mutatta, hogy „a szabályalapú rendszer (Systran) BLEU indexe óriási mértékben alábecsüli a valódi minőséget” (Callison-Burch et al. 2006: 254)

A BLEU módszer néhány hiányossága, amelyek a fenti jelenségekre magyarázatot adhat:

- Szinonimákat és átfogalmazásokat csak akkor talál jónak, ha ezek valamelyik referencia fordításban szerepelnek.
- Minden szó egyenlő súllyal rendelkezik, így a fontos tartalommal bíró részek kihagyása nem jár büntetéssel.
- A rövidebbi büntetés csak egy jobb híján létrehozott mérőszám, hogy kárpótolja a fedés kiszámításának nagyon súlyos hiányát (Callison-Burch et al. 2006: 252).

Boitet et al. viszont az állítja, hogy szabályalapú rendszerek minőségének vizsgálatához is használható a BLEU, de a referenciafordításokként az adott

rendszer nyers gépi fordításainak utószerkesztett változatait kell használni (Boitet 2009: 4).

A fentebb leírt módszerek közül az NIST és a METEOR, valamint Babych és Hartley próbálja javítani ezeket a hiányosságokat, de a különböző rendszerek összehasonlításánál jelentkező problémák, és a következetlenség nem egyedül a BLEU módszerre jellemző, hanem a többi automatikus módszerre is, amelyek szintén a hasonlósági modellt alkalmazzák (Babych és Hartley 2004: , Callison-Burch et al. 2006: 255-256). Gamon et al. az előzőekkel szemben mondat-alapú értékelő módszert ír le, amely kiszűri a hibás mondatokat, referencia fordítások nélkül (Gamon et al. 103-111).

Az itt bemutatott módszereken kívül még sok, hasonló elvű értékelési metrika létezik. Mivel azonban a később bemutatott saját értékelési rendszer az itt bemutatott módszerekre támaszkodik, különös tekintettel a BLEU és a METEOR módszerre, valamint módszerek hasonlósága miatt ezeket nem mutatjuk be részletesen. Csak egy felsorolást szeretnénk róluk adni:

- A ROUGE módszer háromféle különböző illesztést használ (Lin és Och 2004: 605-612).
- A GMT (General Text Matching) módszer unigramokon alapuló F-mértéket számol (Turian et al. 2003: 386-393).
- Albrecht és Hwa a nem tökéletes referenciák lehetőségét vizsgálja (Albrecht és Hwa 2008: 187-190).
- Giménez és Marquez több mérési módszert ötvöz a különböző nyelvi szinteknek megfelelően (Giménez és Marquez 2008: 195-198).
- Popovic és Ney a WER módszer eredményeinek további elemzésével kívánnak a hibákról információt kapni (Popovic és Ney 2007: 48-55).
- Duh a minőségre adott abszolút értékelés helyett rangsort állítanak fel az értékelendő rendszerek között (Duh 2008: 191-194).

## 4. Érthetőség és elfogadhatóság

### 4.1. Bevezetés

Az előző fejezetben szó esett arról, hogyan lehet a gépi fordítás minőségét mérni, hogyan lehet a javulás mértékét felmérni. Ebben a fejezetben bemutatunk egy kutatást, amelynek célja egyrészt felmérni a géppel fordított szövegek érthetőségét és az olvasók szubjektív véleményét A hibaelemzés módszerét Arnold et. al. írja le vázlatosan. Az első lépés a szövegben várhatóan előforduló hibakategóriák listájának felírása. Utána a szöveget elemezve a hibák besorolása következik a kategóriákba.. A hibákat ezek után súlyozni kell. A súlyozásra nem ír egyértelmű módszert, mert a hibaelemzés egyik nehézsége pontosan a súlyozási módszer megállapításában rejlik. A súlyok megállapítására két fő szempontot ad, az első szerint a súlyokat aszerint állítjuk be, hogy a kijavításuk mennyi utószerkesztési művelettel járna. A másik szempont az lehet, hogy a hiba mennyire befolyásolja a szöveg érthetőségét (Arnold et. al. 1994: 166-167). Ezeken kívül egyéb súlyozási módszerek is elképzelhetők, például a hiba mennyire befolyásolja az elfogadhatóságot (azaz mennyire „zavarja”, „idegesíti” az olvasót). Kutatásunkban ezt az irányvonalat követtük a szövegek hibáinak értékelésére.

Az érthetőség mérésére és a hibák feltárására készült egy kérdőív, amelyben az alanyoknak három rövid szöveg elolvasása után a tartalomra vonatkozó kérdésekre kellett válaszolni, valamint egyes kiválasztott rövid szövegrészeket (szavakat, kifejezéseket, mondatokat) kellett megmagyarázniuk. A vizsgálat motivációja annak bizonyítása, hogy az érthetőség és az elfogadhatóság nem azonos, és nem is egyenes arányban állnak egymással. A minőség vizsgálatakor azt gondolnánk, hogy az érthetőségből egyértelműen és egyenes arányban következik az elfogadhatóság is, legalábbis gépi fordítások esetében. A felmérésekből viszont az derül ki, hogy az érthetőség nem elég az elfogadhatósághoz.



## 4.2. A géppel fordított szövegek megértése

### 4.2.1. A szövegértés

Ahogy már többször kihangsúlyoztuk, álláspontunk szerint a gépi fordítás elfogadhatósága a minőség mérésének legfelsőbb mércéje. Az elfogadhatósághoz viszont több tényező járul hozzá. Az egyik ilyen tényező a fentebb említett minőségi szempontok közül szöveg érthetősége. A megértés szintje a géppel készített szövegek esetében egészen más, mint emberek által (anyanyelven) készített szövegeknél.

A szövegek érthetőségéhez a Beaugrande által felsorolt feltételek megléte járul hozzá, legfőképp a kohézió, a koherencia, az informativitás, a szituacionalitás és még az intertextualitás is. Ezeken kívül még a grammatikalitást is hozzá kell tennünk a listához, mint alapfeltétel. (Beaugrande 1983: 4-10). A grammatikalitást általában természetesnek vesszük, de ha a rengeteg, nem anyanyelvi beszélők által írt szövegre, publikált cikke gondolunk, akkor ez a feltétel nem teljesül még emberi szövegek esetében sem. Ha pedig ezeket a szövegeket elfogadjuk, és fel is használjuk – főleg információszerzésre –, akkor a géppel fordított szövegeket sem szabad azért kizárni a szövegek köréből, mert nyelvi hibákat tartalmaznak. A nyelvi hibák viszont nyilvánvalóan rontják a szöveg érthetőségét és így elfogadhatóságát is, a gépi fordítás pedig – széleskörű vizsgálat nélkül is megállapítható – sok nyelvi hibát tartalmaz, így ezzel a problémával foglalkozni kell. A cél az, hogy a hibák számát és súlyosságát lecsökkentsük. A nyelvi hibák a szövegben az érthetőségtől függetlenül is befolyásolják az elfogadhatóságot. A hibás szöveg ugyanis megsérti a nyelvi normákat, ami Heltai szerint az olvasók részéről különböző szankciókkal jár. „A szankciók skálája az enyhe rosszallástól az előítéletig, megbélyegzésig és a diszkriminációig terjed” (Heltai 2004: 422). Ez igaz a gépi fordításra is: ez is magyarázza azt a jelenséget, amelyet szeretnénk bemutatni, hogy a szövegek érthetőségéhez képest az elfogadhatóságuk kisebb.

Intuíció alapján feltételezhetjük, hogy a géppel fordított szövegek nyers formában jelenleg a szövegértési szintek és funkciók közül információ kinyerésére alkalmasak, tehát ez a funkció áll a kérdőíves vizsgálat fókuszában. A felmérésben nem foglalkozunk az alanyok szövegértési képességeivel, hiszen a cél nem a szövegértési képességek javítása, hanem a szövegek minőségének felmérése, a legsúlyosabb hibák kiszűrése, és a szöveg javítása úgy, hogy az átlagos képességű olvasó számára elfogadható legyen. A szövegértést befolyásoló, szövegen kívüli tényezőket a felmérésben ezeket figyelmen kívül hagyjuk, és átlagosnak tekintjük.

Tóth B. a megértésnek három szintjét adja meg, az explicit, az implicit és a kritikai megértést. „Az explicit megértés az elhangzottak szó szerinti felfogása. Az implicit megértés a ki nem mondott, rejtett gondolatok megértését jelenti, alapvetően következtetéseken alapul, s a következtetések levonásában döntő szerepe van a kontextusnak. A kritikai megértés szintjén az egyén értékeli is az elhangzottakat illetve az olvasottakat” (Tóth 2006: 458-459, eredeti forrás Tóth 2002). A három megértési szinttől azt várnánk, hogy egymásra épüljenek, és az alsóbb szintek a felsőbb szintek feltételei legyenek. De a később bemutatott kérdőíves kísérlet tanúságai alapján is az a benyomásunk, hogy az explicit megértés, tehát a szöveg teljes, szó szerinti értelmezése nem elengedhetetlen feltétele a megértés további szintjeinek. A vizsgálatban ugyanis annak ellenére, hogy a mondatok egyes részeit szinte egyáltalán nem sikerült megérteni, a mondat/bekezdés jelentését az olvasók mégis viszonylag jól kikövetkeztették. A relevancia elvével úgy magyaráztuk a jelenséget, hogy az analitikus implikációk megértése nélkül is megkaphatja az olvasó a kontextuális implikátúrákat.

A következő mondat szemlélteti ezt a megállapítást:

(5) Az emberek, akik másfelől **sodródva megtalálják magukat**, megtagadják, hogy elismerjék, hogy ők bármilyen ilyen dolgok és dühös körökben száguldanak körbe miközben túráztatják a motort és összetörik a felszereléseket, a füleik going a **bíbor egy csúnya árnyéka** és a nyelvük, ami megdöbbenően leginkább megromlik.

Az egyes, kivastagított rövid szövegrészek érthetősége ugyanis jóval 0 alatt volt, a teljes mondat tartalmára vonatkozó kérdésre viszont 8% teljesen jól, 74% válasza pedig jó volt, csak kis mértékben hiányos.

Nem azt szeretnénk állítani, hogy ezeknek a szövegeknek a minősége így jó, vagy megfelelő, és hogy megelégedhetünk ezzel a szinttel, hiszen az olvasó a lényegét megértette (vagy kikövetkeztette, igaz, hogy nagyobb erőfeszítés árán), és esetleg a kritikai megértés is megvalósult. Ezek a szövegek egyértelműen nem elfogadható minőségűek, javítani kell rajtuk. A kísérletekkel és az eredményekkel a gépi fordítás hasznosságát szeretnénk bizonyítani, és azt, hogy az első látásra teljesen érthetetlennek, elfogadhatatlannak ítélt szövegek meglepő mértékben mégis érthetőek.

A géppel fordított szöveg olvasásakor két esetet tekinthetünk tökéletes megértésnek. Az egyik, amikor a szöveget ugyanúgy érti meg, mintha az eredeti forrásszöveg lenne az anyanyelve, a másik pedig az a megértési szint, amelyen megértene a szöveget, ha ember fordította volna (feltételezhetően jó minőségben). Ennél többet nem várhatunk, hiszen ha az eredeti szöveget sem értette volna, vagy az emberi fordítást sem, akkor ennél többet nem várhatunk a gépi fordítástól sem. Ha például az eredeti szöveg nem koherens, vagy a kohézió nem megfelelő, az emberi fordítótól még esetleg elvárhatjuk, hogy ezeket megteremti, kijavítja, de a géptől semmiképp sem várhatjuk, hogy „belevarázsolja” ezeket a szövegbe. A gépi és az emberi szövegek különbségei ellenére érdemes megnézni a szövegértés emberi szövegekre kidolgozott funkcióit is. Az olvasás első szintje az értő olvasás, amely a „a szövegben kifejtett gondolatok, információk szó szerinti megértése” (Tóth 2006: 460). A gépi fordítás fejlesztése a megértésnek ezt a szintjét célozza meg, és a következő felmérés is ezt a megértési szintet vizsgálja.

Mindezzel arra szeretnénk rávilágítani, hogy minőség vizsgálatoknál reális elvárásokat kell kitűznünk magunk elé, a gépi fordítások jelenlegi minőségét szem előtt tartva kell összeállítani a vizsgálati anyagot. Ugyanúgy a pragmatikus szemléletet érdemes követni, ahogy ezt a fejlesztők is teszik, mert az eddigi

kutatások tapasztalatai alapján így lehet minőségi javulást elérni. (Ld. korai módszerektől a jelenlegiekig.) Kezdetben tehát érdemes a legsúlyosabb hibákra koncentrálni, és a megértésnek csak az explicit szintjét mérni.

Ezek alapján a kérdőíves felmérés csak az információkinyerés szempontjából vizsgálta a szövegértést. Nem vizsgálta, hogy az olvasó a szöveget szó szerint megérti-e, tehát az értő olvasást sem vizsgálja teljes egészében. Az ideális megértési szint ebből a szempontból a 100%. A jelenlegi felmérés célja egyrészt az, hogy megnézzük, az olvasók milyen arányban értik meg a szövegben található lényegi információt, tehát milyen messze van az érthetőség az ideálistól. Szintén fontos cél annak megállapítása is, hogy ilyen megértési arány mellett mi az olvasók szubjektív véleménye a gépi fordításról, hiszen az elfogadhatóságot ez nagyban befolyásolja.

#### **4.2.2. A vizsgálat menete**

A felmérés kérdőív segítségével történt. Ennek összeállításához először kiválasztottunk néhány (kb. 10) újságcikket a Mirror és a The Times (brit) angol nyelvű internetes napilapokból. A szövegek kiválasztásánál fontos szempont volt, hogy a tartalmuk az olvasók számára újdonság legyen, hogy az esetleg már ismert téma ne befolyásolja a megértést. Praktikus szempontból szintén fontos volt, hogy általános és érdekes témájúak valamint könnyed hangvételűek legyenek. Egyrészt emiatt esett a választás az újságcikkekre. A másik indok az online újságcikkek használatára az, hogy a webfordítás természetes használata (és nagyon jól működő funkciója) a weboldalak, ezen belül a híryanagok lefordítása. A nagyobb értékelési kampányok és felmérések az Európai korpusz mellett híryanagokat használnak, ez is indokolta a szövegtípus kiválasztását (Callison-Burch et al. 2008: 71, Koehn és Monz 2006:104). Az interneten megjelenő cikkeknek fontos a gyors fordítás, a napilapokban megjelenő információ egyik napról a másikra elavul, tehát ilyen esetekben a webfordítás jó eszköz arra, hogy kellő sebességgel a megfelelő információhoz jussunk.

A nyelv és az újságok kiválasztása önkényes volt. Az angol mint forrásnyelv szükségszerű, mivel a MorphoLogic fordítórendszerének az angol és a magyar a forrásnyelve. A célnyelv természetesen a magyar. Mivel arra nem lett volna lehetőség, hogy minden angol nyelvváltozatot vizsgáljunk, a brit angolt választottuk. Későbbi vizsgálatok tárgya lehet a különböző angol nyelvváltozatok összehasonlítása. Azért, hogy a szövegek regiszter szempontjából kicsit heterogénebbek legyenek, a szövegek a kifinomultabb, irodalmibb és bulvár lapokból származnak.

Minden szöveghez az angol eredeti alapján magyarul készültek a szöveg tartalmának legfontosabb pontjaira vonatkozó kérdések. A kérdések általában a cselekményre vonatkoztak, azt feltételezve, hogy a szöveg szereplőit könnyen ki lehet szűrni. Ezután a cikkeket a MetaMorpho szövegfordító program teljesen automatikusan lefordította, majd a téma és a magyar fordítás alapján három szöveg került a kérdőívbe, két hír és egy könnyed hangvételű könyvbemutató egy részlete. A kiválasztás nem volt véletlenszerű, mivel szükség volt egy sok hibát tartalmazó szövegre is a felmérés eredményességének érdekében. Az első két szöveg első benyomás alapján viszonylag kevesebb hibát tartalmaz, könnyebben érthető. A harmadikként választott szöveg FNY-i változata többszörösen összetett mondatokat tartalmaz, és ritkán használt, speciális szókincssel rendelkezik, így az előzetes várakozásoknak megfelelően a fordított CNY-i változata nehezen érhető lett, és a célnak megfelelően sok hibát tartalmaz. Ebből az utolsó, nehezebb szövegből hosszabb-rövidebb részek is bekerültek a kérdőívbe, ezek részleteire vonatkozó kérdésekkel együtt. A szövegrészeket az alanyoknak meg kellett magyarázniuk. A kérdőív elején az alanyokra, a végén pedig a véleményükre vonatkozó kérdések is szerepeltek. A kérdőív alapos kitöltése 30 percet vesz igénybe, de az alanyok szabadon dönthettek, mennyit szólnak rá.

Az alanyok: informatikusok (7 fő), középiskolai tanulók (13 fő), egyetemi hallgatók (42 fő), egyéb (8 fő), összesen 72. fő. A minta sajnos abból a szempontból nem volt reprezentatív, hogy az alanyok nagy része egyetemet/főiskolát végzett

vagy éppen végez. Ezt kissé ellensúlyozza, hogy a felméréskor a motiváció kívülről jön, ami gyengébb, mint a belső motiváció, és így feltételezhetően kevesebb erőfeszítéssel gyengébb teljesítményt eredményez (Deci és Ryan 2000 :55). A kérdőív szövegeit a 11.2 melléklet tartalmazza.

#### 4.2.3. Az érthetőség vizsgálata

Az érthetőség (amely része az olvashatóságnak, gördülékenységnek) méréséhez a kiválasztott három szöveghez három-három kérdést tettünk fel, csak a tartalomra vonatkozóan, méghozzá olyan tartalomra, amelyet az olvasó várhatóan nem ismert előtte. A kérdések nyílt kérdések, a válaszadó tetszőlegesen, saját szavaival megfogalmazhatta a válaszát. Ennek célja egyrészt az, hogy a válaszok közötti finomságokat ki tudjuk szűrni, másrészt hogy a részben kiszűrt információt is meg tudják adni, harmadrészt pedig a zárt típusú kérdések félrevezető eredményt hozhattak volna. A tippelést ebben a felmérésben ugyanis nem engedhetjük meg.

A válaszokat részletes értékelés után hat kategóriába osztottuk, ennél kevesebb nem lett volna elégséges, több pedig nem adott volna jelentősen több információt az elemzéshez. A természetesen adódó *teljesen helyes* és *helytelen*, valamint a *nem tudja* kategóriák között is szükség volt fokozatokra. Elemzéskor részletesen kiszűrtük, hogy a válaszokban mi volt a hiányzó, és mi volt a hozzáadott információ, ezek alapján pedig létrehoztuk a *hiányos*, a *jó és helytelen* és a *nagyrészt helytelen* kategóriákat.

A *hiányos* válaszban minden helyes, de a válasz egy kis része hiányzik. A *jó és helytelen* válasz egy része helyes, viszont a válaszadó félreértett valamit, ezért a válasz egy része helytelen. De ebbe a kategóriába csak azok a válaszok tartoznak, ahol a válasz nagyrészt helyes, csak egészen apró tévedés van benne. (Mivel a válaszokban a tévedések is egy mintát követtek, következetesen meg lehetett állapítani a különbséget az apró vagy nagyobb tévedések között.) A *nagyrészt helytelen* válaszban csak a várt információ egy kis része szerepel, a válasz maradék része vagy helytelen (félreértett) vagy a szövegben nem is szereplő („költött”)

információ. A *helytelen* válasz értelemszerűen az a válasz, amelyben nincs értékelhető rész. A *nem tudja* kategória az, amikor a kérdésre nem volt egyáltalán válasz.

A szövegek érthetőségének mérésére és összehasonlítására egy mérőszámot kell létrehozni, amellyel a különbségek egyértelműen, számszerűen kifejezhetőek. A mérőszám különbséget tesz a nem érthető és a félreérthető szöveg között. A félreérthetőséget eredményező hibát súlyosabbnak ítéljük, mint ha a hiba miatt az adott szövegrész egyáltalán nem érthető.

A kérdőív készítésekor a szövegértésre vonatkozó hipotézisünk az volt, hogy az olvasók az első két szövegre vonatkozó kérdésekre nagyrészt (50% körüli vagy annál magasabb arányban) teljesen jól, vagy jól, de esetleg hiányosan fognak válaszolni. A harmadik szövegben a jó és hiányos válaszok összessége kevesebb lesz, mint 50%. Ezt a hipotézist saját szubjektív véleményre alapoztuk. A második szövegben található egy súlyos referenciális hiba, emiatt várható, hogy az erre vonatkozó kérdés rontani fogja az eredményt.

Az érthetőségre vonatkozó eredmények a következőképpen alakultak:

	Kérdés	Teljes	Hiányos	Jó és helytelen	Nem tudja	Nagy-részt helytelen	Hely-telen	Index
1.	Miért került bajba David Beckham?	30%	24%	12%	6%	22%	6%	0,270
2.	Mi a baj a drótkerítéssel?	28%	24%	4%	30%	4%	10%	0,405
3.	Mi történt Freddie Williamsszel és a lányával?	70%	12%	6%	2%	6%	4%	0,710
4.	Mit vettek el tőlük?	84%	0%	14%	0%	0%	2%	0,820
5.	Mennyit?	8%	2%	26%	0%	60%	4%	-0,460
6.	Mi lett velük?	4%	28%	14%	8%	36%	10%	-0,100
7.	Miről szól az RAC felmérése?	58%	14%	12%	6%	6%	4%	0,615
8.	Mit tesznek a nők, ha eltévednek?	74%	4%	10%	12%	0%	0%	0,800
9.	Mit tesznek/milyenek a férfiak, ha eltévednek?	8%	74%	4%	14%	0%	0%	0,670
	Összesen	41%	20%	11%	9%	12%	7%	0,414

1. táblázat: A kérdésekre adott válaszok relatív megoszlása

#### 4.2.4. Az érthetőségi index meghatározása

Érdekes kérdés annak megállapítása, hogy melyik válasz milyen súllyal számítson a végső értékelésbe. Ahhoz, hogy a szövegekre vonatkozóan egy mértéket kaphassunk, a válaszok típusait számszerűsíteni kell. Ehhez a legkézenfekvőbb módszer az, ha minden választ egy „osztályzatnak” feleltetünk meg, azaz a végső eredménybe ilyen súllyal fog beszámítani. A kérdés az, hogy az osztályzatokat hogyan állapítsuk meg. Ehhez először sorba kell rendezni a válaszok típusait aszerint, hogy a megértés szempontjából melyik jelenti a legjobb és a legrosszabb esetet. Egyértelmű, hogy a skála elején a legjobb érték, a *teljes válasz* fog szerepelni. A következő a *hiányos, de nagyrészt jó* válasz.

A *nem tudja* és a *jó és helytelen* kategória között a következőképpen határoztuk meg a sorrendet: Ha a kérdésre a válaszadó nem tudja a választ, ez egyértelműen



azt jelzi, hogy a szöveg nem érthető. Ha a válaszban van jó rész is, ez azt jelzi, hogy a szöveg legalább részben érthető. Ezek alapján a *nem tudja* választ hátrébb kell sorolni, mint a *részben helyes, részben helytelen* (azaz *jó és helytelen*) választ. A válaszban levő helytelen rész ezzel ellentétben azt jelzi, hogy a szöveg egy része félreérthető, amely a szöveg olvasáskor negatívabb hatással van a szövegértésre, mint a nem értés, hiszen az olvasó abban a hitben marad, hogy megértette a szöveget. Ennek alapján meg kell fontolni, hogy a *nem tudja* válasz ne kapjon esetleg mégis jobb osztályzatot, mint a *jó és helytelen* válasz. A kérdőívek értékelésekor viszont csak azok a válaszok kerültek a *jó és helytelen* kategóriába, amelyekben a válasz döntő többsége jó volt, csak kis tévedés szerepelt benne, így megmaradunk az eredeti felvetésnél, tehát a *nem tudja* válasz nagyobb negatív súllyal fog szerepelni, mint a *jó és helytelen* válasz. A fenti érvelés alapján a *nagyrészt helytelen* kategória pedig még nagyobb negatív súlyt kap, mint a *nem tudja*, hiszen ebben a válaszok azt mutatják, hogy a szöveg nagyobb részben félreérthető volt. A skála végén pedig a *helytelen* válaszok vannak, amelyek azt mutatják, hogy a szövegnek az a része teljesen félreérthető, tehát ezek vezetik legjobban félre az olvasót. A sorrend tehát:

1. Teljes válasz
4. Hiányos válasz
5. Jó és helytelen
6. Nem tudja
7. Nagyrészt helytelen
8. Helytelen

A sorrend eldöntése után pedig súlyokat, vagy osztályzatokat kell rendelni a kategóriákhoz. A gépi fordításban szokásos értékelési módszereket alapul véve a szöveg minőségét praktikus, ha egy 0 és 1 közötti számmal tudjuk megadni. A problémát az jelenti, hogy míg a szövegek minőségének mérésekor a legalacsonyabb érték a rossz minőséget, az egyre magasabb érték pedig értelemszerűen az egyre jobb minőséget jelenti. Ebben a vizsgálatban viszont azok

a részek, amelyek nem érthetőek, és azok, amelyek félreérthetőek egyaránt rossz minőséget jelentenek, és egyik sem elfogadható. Ezek tehát mind 0 vagy ahhoz közeli értéket érdemelnének a skálán, és a minőség szempontjából azt mondhatjuk, hogy nincs köztük különbség, hiszen mindegyik egyértelműen rossz. Viszont az olvasó szempontjából közelítve a félreérthető szöveg sokkal „veszélyesebb”, rosszabb következményekkel jár az olvasó számára. Mert míg a nem érthető szövegnek az olvasóra nincs kontextuális hatása – legalábbis olyan nincs, amely a szöveg tartalmából következne – a félreértett szövegnek van, méghozzá hamis, negatív hatása. A szöveg tehát nemhogy nem lesz releváns, hanem, ha egyáltalán létezik ez a fogalom, akkor negatívan lesz releváns. Ezt a különbséget, a „veszélyesség” mértékét pedig a mérőszámban mindenképp ki kell fejezni.

Ha a mérőszám  $[0;1]$  intervallumba eső pozitív valós szám, akkor a skála közepén fog elhelyezkedni a nem érthető érték, azaz a nem érthető szöveg minősége 0,5 lesz. Ez a többi indexszel (pl. BLEU, vagy a későbbi automatikus vizsgálatban szereplő index) összehasonlítva félreérthető eredményt ad, hiszen ott a 0,5-ös érték már viszonylag jó minőséget jelent. Ezért tehát célszerűbbnek láttuk, hogy az index értéke a  $[-1;+1]$  intervallumban legyen, ahol a 0 jelenti a nem érthető szöveget, a -1 teljesen félreérthetőt, a +1 pedig a teljesen érthetőt. Így valóban teljesül az a követelmény, hogy a  $[0;1]$  intervallumban levő értékek esetén az egyre jobb érték egyre jobb minőséget is jelent. A 0 alatti értékek az elfogadhatatlan minőséget jelentik.

Ezek után létrehoztunk érthetőségi kategóriákat, majd ezek alapján tehát meghatároztuk a különböző kategóriákhoz tartozó számértékeket. A skála nem egyenletesen helyezkedik el az intervallumban, a számok kifejezik azt is, hogy az egyes kategóriák nem egyformán térnek el egymástól. Az érthetőségi kategóriák megfelelnek a válaszok alapján felállított kategóriáknak (amelyek a táblázatban láthatóak). A kategóriáknak megfeleltetett számértékek:

1. Félreérthető (helytelen válasz)    -1

2. Nagyrészt félreérthető (nagyrészt helytelen válasz) -0,75
3. Nem érthető (nem tudja a választ) 0
4. Félig érthető (Részben jó, részben helytelen válasz) 0,25
5. Nagyrészt érthető (Hiányos válasz) 0,75
6. Érthető (Teljes válasz) 1

Minden válasz (azaz a vonatkozó szövegrész) érthetőségi indexét a következőképpen számíthatjuk ki:

$$E = \frac{\sum_{i=1}^6 (n_i \cdot s_i)}{\sum_{i=1}^6 s_i},$$

ahol  $E$  az érthetőségi index,  $i$  a kategória sorszáma,  $n$  az egyes kategóriákban kapott válaszok száma,  $s$  pedig a kategóriára adott számérték. A fenti táblázat utolsó oszlopában ezek az érthetőségi index értékek láthatóak.

Az eredmények alapján úgy ítéltük meg, hogy szükség van a kategóriákhoz tartozó érték-skála finomítására, valamint a jobb olvashatóság kedvéért a 10 új intervallumhoz ugyanezeket a szöveges megnevezéseket rendeltük, jelezve, hogy az intervallum melyik részében helyezkednek el.

- - 1,000 - -0,875 félreérthető
- - 0,875 - -0,750 nagyrészt félreérthető -
- - 0,750 - - 0,375 nagyrészt félreérthető +
- - 0,375 - 0,000 nem érthető -
- 0,000 - 0,125 nem érthető +
- 0,126 - 0,250 félig érthető -
- 0,250 - 0,500 félig érthető +
- 0,500 - 0,750 nagyrészt érthető -
- 0,750 - 0,875 nagyrészt érthető +
- 0,875 - 1,000 érthető

Az indexértékeknek megfelelő szöveges kategóriákat az alábbi táblázat tartalmazza:

Kérdés	Index	Érthetőség
<b>Miért került bajba David Beckham?</b>	0,270	félíg érthető +
<b>Mi a baj a drótkerítéssel?</b>	0,405	félíg érthető +
<b>Mi történt Freddie Williamsszel és a lányával?</b>	0,710	nagyrészt érthető -
<b>Mit vettek el tőlük?</b>	0,820	nagyrészt érthető +
<b>Mennyit?</b>	-0,460	nagyrészt félreérthető +
<b>Mi lett velük?</b>	-0,100	nem érthető -
<b>Miről szól az RAC felmérése?</b>	0,615	nagyrészt érthető -
<b>Mit tesznek a nők, ha eltévednek?</b>	0,800	nagyrészt érthető +
<b>Mit tesznek/milyenek a férfiak, ha eltévednek?</b>	0,670	nagyrészt érthető -
<b>Összesen</b>	0,414	félíg érthető +

2. táblázat: A kérdések az indexekhez tartozó szöveges kategóriákkal

A táblázatban láthatjuk, hogy két szövegrész kapott negatív indexet. A Mennyit? kérdésre a kérdőív 2. szövegében a következő mondat tartalmazta a választ:

(6) Fenygették Mr Williamsöt, Cumnock, Ayrshire, készpénzt adott át - mondott rendőrség által **tíznek éves lenni több ezer**. (GF\_002)

A válaszoknak 64%-a nagyrészt helytelen és helytelen volt, a legtöbb ebből *több ezer*, de többször szerepelt *több millió* is. A *több millió* érdekes, mert a mondatban egyértelműen szerepel az *ezer*, de a szöveg elején szerepel a *millió* szó, tehát itt az olvasó már a szöveg formája helyett inkább tudására és emlékezetére támaszkodik, a feldolgozásra már nem szán akkora erőfeszítést. De ez a jelenség nem befolyásolja az eredményeket, hiszen a várt válasz nem *több ezer*, és nem *több millió*, hanem *több tízezer*. Ez a szövegrész tehát előre láthatóan félreérthető volt, és ezt alátámasztja az érthetőségi index is, amely a *nagyrészt félreérthető +* kategóriát adta. Az, hogy csak nagyrészt félreérthető, az magyarázza, hogy még így is voltak, akik teljesen jól vagy hiányos választ adtak a kérdésre.

A másik negatív indexszel rendelkező szövegrész a *Mi történt Freddie Williamsszel és a lányával?* kérdésre ad választ:

(7) Az hármát vágott az, hogy üveget röpített, és mondtak lenni *\*"traumatised"()*. (GF\_002)

Ennek a szövegrésznek megítéléséhez szubjektív véleményünk már nem elegendő, egyrészt mert a hasonló szövegek olvasása miatti gyakorlat félrevezető,

másrészt az angoltudás segítségével könnyen kitalálható a válasznak legalább egy része. Tehát az ítélethez a kísérlet eredményeire kell támaszkodni. Ez a szövegrész ezek alapján a *nem érthető* – kategóriába esik, ez alátámasztja az elemzés közben szerzett benyomásokat.

#### **4.2.5. Természetes nyelvű szövegek – emberi fordítások**

Az elemzésben az 1-es érthetőségi index jelenti az ideális, tökéletesen érthető szöveget. De ezt az értéket az emberi szövegek sem feltétlenül érik el, tehát a kapott eredményeket nem ehhez az értékhez kell hasonlítani. Hitelesebb a vizsgálat akkor, ha azt is megnézzük, hogy ugyanezeknek a szövegeknek az emberi fordításának mekkora az érthetőségi indexe. Az ideálisnak érthetőségi indexnek így azt tekinthetjük, amennyire egy átlagos olvasó az ember által lefordított szöveget megérti. Ahhoz, hogy meg tudjuk állapítani, az emberi fordítások ugyanezzel az értékelési módszerrel mennyire érthetőek, egy másik felmérést is végeztünk. Ennek a felmérésnek az eredményét tekinthetjük teljes megértésnek.

A felmérésben ugyanazok a szövegek szerepeltek, amelyek a korábbi kérdőívben, ugyanazokkal a kérdésekkel. A kiértékelés ugyanazokkal a kategóriákkal történt. Azt nem várhatjuk ember által létrehozott szövegeknél sem, hogy minden olvasó teljes mértékben megérti, hiszen a megértésnek több szintje van – amelyeket itt nem vettünk figyelembe –, és a szövegértési képességek nem egyformák, valamint a külső motiváció miatt az olvasótól kívánt feldolgozási erőfeszítés esetleg nagyobb mértékű, mint amit a remélt kontextuális hatás kárpótolhatna. Ezek a szövegek viszont saját vélemény alapján elég egyszerűek, rövidek és nem igényelnek háttértudást, tehát a kérdésekre adott teljes válaszok aránya itt közel 100% kell, hogy legyen, az index értékétől pedig 0,95 vagy nagyobb értéket várunk.

Az így kapott eredmények:

	Kérdés	Teljes	Hiányos	Nem tudja	Jó és helytelen	Nagyrészt helytelen	Helytelen	Index
1.	Miért került bajba David Beckham?	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	1,000
2.	Mi a baj a drótkerítéssel?	93,6%	6,4%	0,0%	0,0%	0,0%	0,0%	0,984
3.	Mi történt Freddie Williamsszel és a lányával?	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	1,000
4.	Mit vettek el tőlük?	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	1,000
5.	Mennyit?	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	1,000
6.	Mi lett velük?	87,2%	8,5%	2,1%	0,0%	2,1%	0,0%	0,926
7.	Miről szól az RAC felmérése?	97,9%	0,0%	0,0%	0,0%	2,1%	0,0%	0,963
8.	Mit tesznek a nők, ha eltévednek?	95,7%	2,1%	0,0%	0,0%	2,1%	0,0%	0,957
9.	Mit tesznek/milyenek a férfiak, ha eltévednek?	74,5%	25,5%	0,0%	0,0%	0,0%	0,0%	0,936
	Összesen	94,3%	4,7%	0,2%	0,0%	0,7%	0,0%	0,974

3. táblázat: Az emberi fordításoknál a válaszok relatív gyakorisága

Láthatjuk, hogy az index itt sem 1, tehát valóban nem 100%-os a megértés, de minden kérdés még az érthető kategóriába tartozik, a teljes érthetőségi index pedig valóban 0,95 felett van.

#### 4.2.6. Az eredmények elemzése

A géppel fordított szövegek összesített indexe tehát 0,56 értékkel tér el az emberi fordítások érthetőségétől, ami 28%-ot jelent a pontozási rendszerünk alapján. De ha a szöveget csak a 0 értéktől tekintjük bármennyire is jónak vagy elfogadhatónak, akkor az érthetőségi indexérték alapján azt mondhatjuk, hogy a géppel fordított szövegek nagyjából félig érhetőek. A kezdeti feltételezés részben beigazolódott, mivel az első hat kérdésből csak kettőre nem érte el a helyes vagy hiányos válaszok aránya az 50%-ot. Az 5. kérdésnél a referenciális hiba valóban nagyban befolyásolta az eredményt: a jó és hiányos válaszok aránya itt csak 10%.

Érdekes viszont, hogy a harmadik, ránézésre nehezen érthető és olvasható szöveg esetében a kezdeti feltételezéssel ellentétben két kérdésre nagy arányban teljesen jól válaszoltak (59% és 74%), ha pedig összeadjuk a teljes és hiányos válaszokat, akkor az eredmény 72% és 78%. Az utolsó kérdésre ugyan csak 8% adott teljes választ, de a hiányos válaszokkal együtt a jó válaszok aránya már 82%. Érdekes, hogy erre a kérdésre az emberi fordításoknál is 25% volt a hiányos válaszok aránya, amit soknak ítéltünk. Ennek fényében nem csodálkozhatunk, hogy a géppel fordított szövegeknél ennyire kevesen adtak teljes választ.

Az utolsó szöveg esetében kapott meglepően jó eredményre a később következő, részletes hibaelemzési részben keresünk választ. Az eredmények alapján a kis mértékű gördülékenységi és olvashatóság, a rengeteg hiba ellenére az olvasók a szöveg fő témáját és a „cselekményét” ki tudták szűrni.

Az itt bemutatott elemzési módszerrel létrehoztunk egy indexet, egy mutatót, amely a géppel fordított szövegek érthetőségét méri. Emberi – nem nyelvtanulói – fordítások érthetőségének mérésére véleményünk szerint nem alkalmas, hiszen az emberi szövegek megértésében a különbségek ehhez képest csak árnyalatnyiak, amelyet ez az index nem tud kezelni.

#### **4.2.7. A gépi fordítás elvárt minősége**

Az elemzés utáni eredmények felhasználásával készült egy szintén kérdőíves felmérés, amely a minimálisan utószerkesztett szövegek érthetőségét vizsgálja. A minimális utószerkesztés kategóriáit a fenti részletes hibaelemzés adta, a vizsgálat eredményeit pedig részletesen a 8. fejezet mutatja be. Ebben a kérdőívben azt is megkérdeztük, hogy az alanyok milyen minőséget várnak el a gépi fordítástól. A kérdőívet kitöltők egyike sem jártas a gépi fordításban, sőt, a fordításban sem, de az internetet – egyetemi hallgatók lévén – nagy valószínűséggel az átlagnál gyakrabban használják. Ők lehetnek tehát a gépi webfordítás egyik célcsoportja. A kérdés pontosan így szólt, és három válaszlehetőség közül lehetett választani:

A gépi fordítás azonnal a rendelkezésre áll, nagyon gyors és olcsó. Ennek tudatában Ön milyen minőséget vár el a gépi fordítástól?

- a) *bármí jó, lényeg, hogy megértsem*
- b) *legyen jól olvasható, de nem baj, ha a stílusa furcsa, és nem hasonlít az emberi fordításhoz*
- c) *legyen teljesen olyan, mint az emberi fordítás*

Előzetes tapasztalatok, és a korábbi kérdőívre adott szubjektív válaszok alapján azt vártuk, hogy a válaszadók nagy része, nagyjából 65% a második lehetőséget fogja választani, a c) válasz kapja a második legtöbb szavazatot, és az a) válasz lesz az utolsó lehetőség a listában. A korábbi kérdőív alapján ugyanis a válaszadók elvárásai nagyok voltak, véleményük pedig igen kritikus. Meglepő módon viszont az eredmények a következőképpen alakultak:

Minőség	Válaszok száma	Válaszok aránya
Bármí jó, lényeg, hogy megértsem	13	27,66%
Legyen jól olvasható, de nem baj, ha a stílusa furcsa, és nem hasonlít az emberi fordításhoz	25	53,19%
Legyen teljesen olyan, mint az emberi fordítás	7	14,89%
Nincs válasz	2	4,26%

**4. táblázat: A gépi fordítástól elvárt minőség**

A legmagasabb arányban valóban a második választ választották, de ami a várakozásoknak nem felelt meg, az az a) válaszok nagy aránya. Tehát ebben a közönségben viszonylag kevesen várnak olyan fordítást a géptől, mint az emberi szöveg. A gépi fordítás szempontjából pozitívak a válaszok, mert ez azt mutatja, hogy a nagy többség elvárásai nem irreálisak, és az előnyök figyelembevételével áldozatokra, nagyobb erőfeszítésre is hajlandóak.

Ezek a válaszok érdekesek ahhoz képest, amit az első kérdőíves felmérés eredményeként kaptunk. Az első kérdőívben egyrészt azt kérdeztük meg, hogy hogyan változott az olvasó véleménye a gépi fordításról. A kérdés feltevésekor azt



feltételeztük, hogy az olvasó elvárása az, hogy a szövegből a fontos információt ki tudja szűrni, és hogy a szöveget meg tudja érteni. Ez alapján és mivel a szövegeket félig megértették, azt vártuk volna, hogy azoknak az aránya, akiknek romlott a véleménye, nagyjából 50% körül lesz. Ehhez képest 60% válaszolta, hogy romlott a véleménye, 2%-nak csak kicsit romlott. 10%-nak nem változott, de eleve rossz volt a véleménye. Ez összesen 72%, úgy ítéljük, hogy számukra a gépi fordítás minősége nem elfogadható. A maradék 28%-nak a fele egyáltalán nem válaszolt, a másik fele pedig azt válaszolta, hogy a véleménye nem változott. Egyetlen olyan válasz sem volt, amelyben javuló véleményt adtak volna meg.

Ebből arra következtetünk, hogy a kezdeti feltételezés nem volt helyes, az olvasók elvárásai nem a jelenlegi helyzetet tükrözik, valószínűleg nagy részük nem találkozott még gépi fordítással. Ez viszont ellentmond az előző felmérés eredményeivel, amelyben pedig realisztikus elvárásokat adtak meg. Ez a felmérés 2006-ban készült, az előző, az elvárásokat mérő pedig 2011-ben, tehát lehetséges, hogy a különbség a közben eltelt időből is fakad. A különbség magyarázata az is lehet, hogy az olvasók számára ez a szöveg egyáltalán nem tartozik az olvasható kategóriába, és úgy ítélik meg, hogy egyáltalán nem értik meg. Ezt pedig megcáfolja a korábban bemutatott érthetőségi index, tehát sokkal többet megértenek a szövegekből, mint amennyit ők maguk is gondolnak. A szövegek elfogadhatósága és érthetősége tehát a feltételezettől egymástól kevésbé függő kategória, és az érthetőség nem elég az elfogadhatósághoz. Az elfogadhatóság tehát nagyban függ egyéb tényezőktől is, pedig a gépi fordítás esetében a mesterséges nyelv sajátosságait meg kellene szokni, és az elfogadhatósághoz elegendőnek kellene lennie az érthetőségnek.

Szintén fontos eredmény, hogy a válaszadók 74%-a többszöri elolvasás után jobban értette a szöveget. Ez egyrészt azt jelenti, hogy a feldolgozási erőfeszítés növelésével növekszik a kontextuális hatás, tehát az érthetőséget már az is javítja, ha az erőfeszítést növeljük. Ha a szövegek egyáltalán nem lennének érthetőek,

akkor az erőfeszítés mértéke nem befolyásolná a kontextuális hatást. Másrészt pedig azt jelenti, hogy a szövegek nehéz megértését szokatlanságuk is okozza.

#### **4.2.8. További szövegek érthetősége**

A részletes vizsgálatok befejezése után az eredmények igazolásához elvégeztünk egy újabb felmérést ugyanezzel a módszertannal. 30 fő töltött ki hetente 1-1 kérdőívet 1 szöveggel, 6 héten keresztül. Ezek a szövegek legalább kétszer olyan hosszúak voltak, mint az eredeti kérdőíven szereplő szövegek, és ugyanúgy nyitott kérdések szerepeltek a szövegek végén. Ezeknek a szövegeknek az átlagos érthetőségi indexe jelentősen jobb lett, mint az első kérdőív szövegei: 0,663. Ez az eredmény is igazolja, hogy a szövegek érthetősége jobb az elfogadhatóságuknál.

#### **4.3. Az eredmények összefoglalása**

A szövegek érthetőségéről az indexértékek alapján azt mondhatjuk, hogy ezekben az esetekben a gépi fordítás alkalmas az értő olvasásra, a szöveg lényegének kiszűrésére. Az egyes részletek elemzése után pedig feltételezésünk az, hogy ha a legsúlyosabb, a megértést legjobban akadályozó hibákat kijavítjuk, akkor a szöveg érthetősége ugrásszerűen meg fog nőni, az érthetőségi indexértéket tekintve akár 4 tizeddel is. A legsúlyosabb hibáknak azokat tekintjük, amelyeknek a – 7. fejezetben leírt – javítási indexe 0 vagy negatív. Az erre vonatkozó felmérést a 8. fejezetben mutatjuk be.

Az érthetőség mérésére készült egy érthetőségi index, amelynek értéke  $[-1;+1]$  között van. A  $[-1;0[$  közötti érték jelzi, hogy a szöveg mennyire félreérthető, a  $[0;1]$  közötti érték pedig azt, hogy a szöveg mennyire érthető. Az érthetőségi index alkalmazható géppel fordított szövegek, utószerkesztett szövegek és emberi szövegek érthetőségének mérésére. A felmérésekből kiderült, hogy a géppel fordított szövegek érthetősége az index alapján 0,414, az emberi szövegek érthetősége pedig 0,974. A 0,414-es érthetőségi index 28%-kal alacsonyabb megértési szintet jelent az emberi szövegek érthetőségénél. A későbbi, több

szöveget vizsgáló felmérés eredményében a szövegek érthetősége még jobbnak bizonyult, az érthetőségi index 0,663 volt. Az első érték a félig érthető kategória felső részébe esik, ez utóbbi pedig a nagyrészt érthető kategória alsó felébe, ami a szövegek jellegét tekintve jó érthetőségnek számít.

Az érthetőségi eredményekhez képest az olvasók 72%-ának az elvárásainak nem felel meg a gépi fordítás, azaz számukra elfogadhatatlan. A két eredmény azt jelenti, hogy az érthetőség és az elfogadhatóság nem azonos kategória, és egymással nem állnak egyenes arányban, pedig a gépi fordítás kutatásának és használhatóságának szempontjából ez lenne a kívánatos.

## 5. Az elfogadhatóság mérése

### 5.1. Bevezetés

A 3. fejezetben részletesen bemutattuk a fordítás minőségét mérő módszereket, különös tekintettel az automatikus módszerekre. Az előző, 4. fejezetben egy kézi, kérdőíves módszert írtunk le az érthetőség felmérésére, és bemutattuk, hogy az érthetőség és az elfogadhatóság nem azonos.

A jelenlegi fejezet egy automatikus mérési módszert mutat be, amelynek fő célja a gépi fordítás elfogadhatóságának és olvashatóságának vizsgálata. Amigó et al. végzett kutatásokat annak megállapítására, hogy az elfogadhatóság és az emberi szövegekhez való hasonlóság mennyire egymást fedő kategória. Ők azt bizonyították felmérésükkel, hogy elfogadhatóság tágabb kategória, tehát nem minden szöveg volt az emberéhez hasonló, amelyet az értékelők elfogadhatónak tartottak. De mivel az elfogadhatóság a tágabb kategória, a másik oldalról az emberéhez való hasonlóság garantálja az elfogadhatóságot (Amigó et al. 2006: 22-23).

Hovy et al. alapján (Hovy et al. 2002c: 1-7) minden automatikus mérési módszernek a következő tulajdonságokkal kell rendelkeznie:

- A mérési eredménynek legyen felső határa (1), amelyet a minőség csak akkor érhet el, ha tökéletes.
- Az eredménynek legyen alsó határa (0), amely a lehető legrosszabb minőséget jelenti. Ezt az értéket csak akkor érheti el, ha a minőség abszolút rossz.
- Monotonitás: A eredményeknek monotonnak kell lenniük abban az értelemben, hogy ha A rendszer minősége jobb, mint B rendszeré, akkor az A rendszer eredményének is nagyobbabbnak kell lennie.

Ezekon kívül az értékelési módszerrel kapcsolatos követelmények:

- könnyű meghatározni, világos, magától értetődő

- az emberi ítélettel jól korrelál minden körülmények között, valamint minden műfajban és területen
- a lehető legkisebb különbséget mutat értékelők (emberi értékelés esetében) vagy ekvivalens bemenetek között
- olcsó elkészíteni és használni
- ha lehet, automatikus legyen

A motiváció a módszer kidolgozására az előzőleg bemutatott kérdőíves kutatás volt, amelyből kiderült, hogy bár az alanyok a géppel fordított újságcikkek nagy részét megértik, a vélemények a szövegekről mégis negatívak voltak, tehát az érthetőségük ellenére a szövegek nem elfogadhatóak. Az értékelés tehát annak megállapítására is készült, hogy a gépi produkció mennyiben hasonlít az emberi nyelvhasználatához és a minősítés is pontosan ezen a mértéken alapul. Így Amigó et. alapján ha a gépi fordítás nagyban hasonlít az emberek által használt nyelvvezetethez, a szövegeknek elfogadhatóaknak kellene lenniük (Amigó et al. 2006: 22-23).

A módszer csak az elfogadhatóságot és közvetetten a koherenciát vizsgálja, és nem vizsgálja, hogy a fordítás mennyiben felel meg a FNY-i szövegnek. Az elfogadhatóság szubjektív, kézi módszerrel is mérhető, az automatikus vizsgálat viszont lehetőséget ad az objektív mérésre, több szöveg rövidebb idő alatt feldolgozható, valamint a kis különbségeket is ki tudja szűrni. Az értékelés célja jelenleg is egyrészt az elfogadhatóság megállapítása az emberi szövegekhez képest, valamint a javulás mértékének mérése utószerkesztés (vagy akár a rendszer módosítása) után.

A módszer korpusz alapú, azaz az értékelendő szöveg részeinek előfordulásait vizsgálja egy korpuszban. A korpusz az interneten található magyar szövegek összessége. A korpusz megfelelő, mert nyitott, tehát bárki hozzájárulhat, nem korlátozódik egyfajta - pl. irodalmi - nyelvhasználatra, és kellően nagy. Fontos feltételezés szintén, hogy az itt közzétett szövegek szerzői saját maguk és közönségük számára elfogadhatónak tartják szerzeményeiket, még ha ezek nem is

helyesek, ideálisak, esetleg a közönség egy másik részének még csak nem is elfogadhatóak.

## 5.2. Az értékelés módszere

A kidolgozott értékelési módszer teljes mértékben automatikus, azaz a vizsgálat egy számítógépes programmal történik. A program a vizsgálat végén egy 0 és 1 közé eső számot ad eredményként, amely a szöveg elfogadhatóságát jelzi. Az értékelés részben a BLEU és az NIST módszerére támaszkodik.

Az értékelés a következőképpen működik: Először a program az értékelendő magyar nyelvű szöveg mondataiból  $n$ -gramokat készít. Az  $n$ -gram egy  $n$  db szóból álló szószorozat, tehát először vesszük a mondat szavait, ezek az 1-gramok, utána veszünk minden egyes szót az utána levő szóval együtt, ezek a 2-gramok, és így tovább. Egy 10 szóból álló mondatból értelemszerűen 10 db 1-gram, 9db 2-gram (szópár), 8 db 3-gram, stb. lesz. A program minden  $n$ -gramra egyesével rákeres az interneten különböző keresőkkel, és kiszűri a találatok számát minden egyes sorozatra.

A találati szám ezután súlyt kap, mivel a hosszabb sorozatok megtalálása nagyobb jelentőségű, mint egyetlen szó vagy szópár előfordulása. A hosszabb sorozat előfordulása ugyanis azt jelenti, hogy a szerzők a szószorozatokat pontosan ebben a formában használják, azaz az illető  $n$ -gram így elfogadható. A hosszabb sorozatra kapott találat nagyobb gördülékenységet is sugall, tehát az  $n$ -gram annál nagyobb súlyt kap, minél hosszabb (minél nagyobb az  $n$ ). A gördülékenység nagyban hozzájárul a szöveg elfogadhatóságához.

A használt keresők a Google<sup>2</sup>, a Bing<sup>3</sup> és a Lycos<sup>4</sup>. Ezekre a keresőkre kísérletezés során esett a választás, mivel ezek nagyjából hasonló eredményt adnak, míg a többi (pl. MSN, Tangó, Kurzor) találati számai jelentősen eltérnek, a

---

<sup>2</sup> [www.google.hu](http://www.google.hu)

<sup>3</sup> [www.bing.com](http://www.bing.com)

<sup>4</sup> [www.lycos.com](http://www.lycos.com)

nyelvi intelligenciát használó PolyMeta pedig már a használhatatlanságig lelassította a programot. Mindhárom keresőben van lehetőség a nyelv beállítására, ami rendkívül fontos. Ha a kereső minden nyelven keresi az n-gramokat, rövidebb n-gramoknál ( $n = 1, 2$ ) az eredmény félrevezető lehet.

A n-gramoknál az  $n$  megválasztásához szintén előzetes kísérletek folytak. A keresések alapján az derült ki, hogy míg 5 szóból álló sorozatokra vannak találatok, szóhatosokat szinte soha nem talál egy kereső sem, így az  $n$  maximális értéke az értékelésnél kezdetben 5.

Biztonsági okokból – a megbízhatóság érdekében – 1 és 2 betűből álló szavakat a program nem keres, a hárombetűs szavaknak pedig nagy részét kizárja, mivel ezek lehetnek rövidítések, vagy más nyelvű szövegekben is nagy számban előfordulhatnak. (Bár a keresőkben a nyelvet be lehet állítani, egyes szavakra mégis adnak találatokat más nyelvű oldalakról is.) A találatok kiszűrése után a program ezeket összesíti és kiszámítja az eredményeket.

### 5.3. Teszteléshez használt szövegek

Ahhoz, hogy a módszert megfelelően tudjuk hangolni, és az alkalmazhatóságát, megbízhatóságát meg tudjuk állapítani, szükségünk van referencia-eredményekre „jó” és „rossz” szövegekről is. A géppel fordított szövegeknek mindenképp ezeknek az eredményei között kell elhelyezkedniük ahhoz, hogy az emberi ítélettel megegyezzen az eredmény. Az értékelési módszer megbízhatóságának mindenképpen szükséges - de nem elégséges – feltétele, hogy a „jó” szövegek eredménye közel legyen 1-hez, a rossz szövegeké pedig 0-hoz (Hovy et al. 2002c: 3). A hangolás során felhasználjuk a korábbi kérdőíves kutatásban szereplő szövegeket, hiszen ezek érthetőségéről és kis mértékben az elfogadhatóságukról is van információnk. Ezeken kívül a hangoláskor szintén támaszkodtunk az egyéni szubjektív véleményünkre. Ha a programot megfelelően hangoltuk, akkor alkalmas lesz új szövegek minősítésére.

A teszteléshez használt „jó” szövegek internetes újságokból vett cikkek – mivel a géppel fordított szövegek is ilyenek –, a „rossz” szövegek pedig mesterségesen kreált – szövegnek talán nem is nevezhető, – de annak látszó – karaktersorozatokat. Több ilyen szöveg is készült: az egyik egy cikk szavainak permutációja, egy másik úgy készült, hogy egy cikk szavai közé véletlenszerű (csak betűkből álló) karakterláncok kerültek, a harmadik pedig szószzerű karakterláncok sorozata.

## 5.4. Az eredmények kiszámítása

A program tehát a mondatok  $n$ -gramjait keresi meg az interneten, a találatokból pedig kiszámítja a szöveg minőségét jelző mérőszámot. A szövegre kapott eredmény az egyes  $n$ -gramokra kapott találati számok súlyozott átlaga, egy  $[0,1]$  intervallumba eső szám, mely a minőségét – elfogadhatóság és gördülékenység – jelzi.

Az eredmény meghatározásában a nehézséget az jelenti, hogy az  $n$ -gramokra kapott találati szám felső korlátja tulajdonképpen meghatározhatatlan, míg az eredménynek egy adott intervallumba kell esnie, így a találati számokat le kell képezni a  $[0,1]$  intervallumra. (Az intervallum 0 és 1 határa önkényesen választott, a szakirodalomban általánosan elfogadott és használt intervallum.)

A leképezéshez kétféle módszert próbáltunk ki. Az első módszer nem veszi figyelembe a találati arányokat, csak azt a tényt, hogy az  $n$ -gramra volt találat vagy nem. Így tehát a megtalált oldalak számától függetlenül az  $n$ -gram 1 vagy 0 értéket kap, majd ez az érték kap súlyt az  $n$ -gram hossza alapján.  $Q_1$  jelöli a végső eredményt,  $n$  a sorozatokban levő szavak számát (az  $n$ -gram hosszát),  $r_s$  az  $n$ -gram találati számát, a  $w_n$  pedig az  $n$ -gramok súlyait. Az  $n$  maximális értéke 5.  $m$  a mondat szavainak száma. Így

$$Q_1 = \frac{\sum_{n=1}^5 \sum_{s=1}^{m-n+1} (r_s \cdot w_n)}{\sum_{n=1}^5 \sum_{s=1}^{m-n+1} w_n}$$



Egy mondat  $n$ -gramjainak száma összesen  $S$ :

$$S = \frac{m(m+1)}{2},$$

ami pontosan  $S$  keresést jelent, így a program bonyolultsága polinomiális,  $O(m^2)$ . A két módszeren belül is kétféle elemzés történt, a  $Q_{1A}$  eredmény a szavakat nem veszi figyelembe, tehát  $n = 2, \dots, 5$ , a  $Q_{1B}$  eredményben az  $n = 1, \dots, 5$ .

A második módszer az eredmények leképezésére figyelembe veszi a találati gyakoriságot is. Mivel a minősítés így egy tetszőleges pozitív valós szám lehet, minden  $n$ -re és minden keresőprogramhoz egy-egy gyakorisági számot kellett megállapítani, majd a találatok számát ezzel a számmal kell elosztani azért, hogy az eredmény végül a  $[0,1]$  intervallumba essen. A leképezéshez felmerült valamilyen szigorúan monoton,  $N \rightarrow [0,1]$  lineárishoz közelítő függvény használata is, de végül a gyakorisági számok is megfelelőnek bizonyultak. A gyakorisági szám  $F_{nj}$ :  $n$  az  $n$ -gram hossza,  $j$  pedig a keresőprogram sorszáma ( $n = 1, \dots, 5$ ,  $j = 0, \dots, 2$ ). Az  $F_{nj}$  egy olyan szám, amely kifejezi, hogy egy  $n$  hosszúságú sorozat találatainak számát mikor tekinthetjük már kellően nagyoknak. Minden  $n$ -gram hosszúsághoz és minden keresőhöz egy-egy gyakorisági szám tartozik. Így

$$Q_2 = \frac{\sum_{n=1}^S \sum_{s=1}^{m-n+1} \left( \frac{r_s \cdot w_n}{F_{nj}} \right)}{\sum_{n=1}^5 \sum_{s=1}^{m-n+1} w_n} \quad j = s \bmod 3,$$

ahol  $m$  a mondat szavainak száma,  $n$  az  $n$ -gram hossza,  $r_s$  az  $n$ -gramra kapott találatok száma,  $w_n$  az  $n$ -gram hossza alapján megállapított súly. Ebben az esetben is a  $Q_{2A}$  képletben  $n = 2, \dots, 5$ ,  $Q_{2A}$  képletben pedig  $n = 1, \dots, 5$ .

Ha  $\frac{r_s \cdot w_n}{F_{nj}}$  nagyobb, mint 1, az  $n$ -gram eredménye 1 lesz, hiszen ekkor az  $n$ -gram kellően gyakorinak mondható ahhoz, hogy elfogadható és gyakran használt legyen.

A gyakorisági számok meghatározása empirikus módszerrel történt: számos keresés és az eredmények tesztelése után kiderült, hogy ha egy szó 10000-szer előfordul, egy szópár 200-szor, egy szóhármás 100-szor, egy szónégyes 10-szer, egy szóötös pedig akár csak egyszer is, az már gyakorinak és széles körben használnak tekinthető. Ezeket a számokat a kissé a különböző keresőkhöz kellett alkalmazni, és így megfelelőnek bizonyultak.

A kísérleti eredmények közül néhányat az alábbi táblázat mutat be. G-vel jelöltük a jó, emberi szöveget, B-vel a mesterségesen létrehozott rossz szövegeket, MT-vel pedig a géppel fordított szövegeket.

<b>Teszt 1</b>	nagy súlyok (0.1, 10, 100, 1000, 10000, 10000), nagy gyakorisági számok			
	<b>Q<sub>1A</sub></b>	<b>Q<sub>1B</sub></b>	<b>Q<sub>2A</sub></b>	<b>Q<sub>2B</sub></b>
<b>G</b>	0.95	0.999	0.56	0.38
<b>B1</b>	0.51	0.11	0.016	0.01
<b>B2</b>	0	0.0064	0	0.0000024
<b>MT1</b>	0.98	0.97	0.57	0.0026
<b>MT2</b>	0.87	0.92	0.035	0.032
<b>Teszt 2</b>	nagy súlyok (0.1, 10, 100, 1000, 10000, 10000), kisebb gyakorisági számok			
	<b>Q<sub>1A</sub></b>	<b>Q<sub>1B</sub></b>	<b>Q<sub>2A</sub></b>	<b>Q<sub>2B</sub></b>
<b>G</b>	0.95	0.999	0.91	0.91
<b>B1</b>	0.51	0.11	0.36	0.051
<b>B2</b>	0	0.0064	0	0.0029
<b>MT1</b>	0.93	0.97	0.56	0.47
<b>MT2</b>	0.87	0.92	0.66	0.3
<b>Teszt 3</b>	végső számok: kisebb súlyok (0.1,2,5,10,50,100), kisebb gyakorisági számok (10000,200,100,10,1)			
	<b>Q<sub>1A</sub></b>	<b>Q<sub>1B</sub></b>	<b>Q<sub>2A</sub></b>	<b>Q<sub>2B</sub></b>
<b>G</b>	0.98	0.98	0.91	0.92
<b>B1</b>	0.08	0.052	0.079	0.012
<b>B2</b>	0	0.061	0	0.0047
<b>MT1</b>	0.75	0.79	0.49	0.45
<b>MT2</b>	0.36	0.33	0.26	0.25

5. táblázat: Az előzetes tesztek eredményei

A súlyok meghatározása szintén empirikus módszerrel történt. A súlyokat és a gyakorisági számokat a következő táblázat foglalja össze:

<b>n</b>	1	2	3	4	5
<b>F<sub>n</sub></b>	10000	200	100	10	1
<b>w<sub>n</sub></b>	0,1	1	5	10	50

**6. táblázat: A végső súlyok és gyakorisági számok**

Ha a teljes mondatot sikerült megtalálni, ez 50-szeres súlyt kap, hosszára való tekintet nélkül, mivel ez azt jelenti, hogy a teljes mondat elfogadható. Ha egy n-gramot ( $n = 3, \dots, 5$ ) nem sikerült megtalálni, ekkor ezek súlya módosult 1-re, ha pedig egy szóra nem volt találat, ez 100-szoros súlyt kapott.

Nézzünk egy példát a Q<sub>2B</sub> módszerre:

<b>N-gram</b>	<b>Találatok száma</b>	<b>Súlyozott eredmény</b>
Vajon	686000	1
lesz	12500000	1
eredmény	756000	1
Vajon mi	53900	2
mi lesz	1180000	2
lesz az	1820000	2
az eredmény	234000	2
Vajon mi lesz	1470	5
mi lesz az	51500	5
lesz az eredmény	8090	5
Vajon mi lesz az	164	10
mi lesz az eredmény	390	10
Vajon mi lesz az eredmény	15	100
<b>Végeredmény:</b>		<b>1</b>

**7. táblázat: Példa a Q<sub>2B</sub> módszerrel kapott eredményre.**

### 5.4.1. A változatok összehasonlítása

A végső értékelési módszer kialakításához a fenti négy változat közül egyet kell választani, és ha kell, még tovább finomítani. Nézzük meg, milyen eredményeket kapunk ezzel négy változattal, a fent említett jó és rossz tesztszövegekkel, valamint több gépi fordítással:

	Q <sub>1A</sub>	Q <sub>1B</sub>	Q <sub>2A</sub>	Q <sub>2B</sub>
„Jó” szövegek	0,87	0,87	0,63	0,65
„Rossz” szövegek	0,11	0,156	0,089	0,085
Gépi fordítás	0,54	0,58	0,40	0,40

**8. táblázat: A különböző szövegek eredményeinek összehasonlítása**

A táblázatban láthatjuk a különböző módszerekkel lefolytatott vizsgálatok átlagait. Ebből először is kiderül, hogy a gépi fordítás eredményei a szubjektív emberi értékelés után az elvárt pozícióban vannak a jó és a rossz szövegek között. Az is látszik, hogy az A és B módszerek között jelentéktelen eltérés van, amit valószínűleg az okoz, hogy a szavakra ( $n = 1$ ) kapott találatok igen kicsi súlyt kapnak. Mégis fontosak a szavakra kapott találatok, mert a géppel fordított szövegekben gyakran nem létező összetett szavak szerepelnek.

Az 1-es és a 2-es módszer eredményei között jelentős eltérés van, a 2-es módszerben a 0-hoz sokkal közelebb esnek az eredmények minden szöveg esetében. Az értékelési módszernek viszont olyannak kellene lennie, hogy a jó szövegre kapott minősítésnek a felső értékhez közel kell esnie. Ebben az esetben a 0,63 és 0,65 értéket nem ítélnék kellően közelinek az intervallum felső határához, azaz 1-hez. Emiatt a Q<sub>2</sub>-es módszert kizárjuk.

Az eredmények alapján tehát úgy ítélnék meg, hogy a Q<sub>1</sub>-es módszer adja a legmegbízhatóbb eredményeket, tehát ez lesz az, amelyet végső értékelési módszernek választunk.

#### **5.4.2. Az értékelési módszer finomhangolása**

Végső értékelési módszernek tehát a Q<sub>1A</sub> módszert választottuk, de a teszteléskor kis mennyiségű szöveget használtunk. A géppel fordított szövegek minősége sem homogén, de azt nem teszteltük, hogy ezek között hogyan tesz különbséget a program. Ezért további tesztekre van szükség, és ha ezek a tesztek indokolják, akkor a módszert tovább kell hangolni.

A tesztek alapján azt láttuk, hogy a program nem mindig következetesen tesz különbséget a szövegek között. Ezért megvizsgáltuk, hogy mekkora különbség

van a szövegeknél (emberi és gépi fordítás) az egy  $n$ -gramra kapott átlagos találati számokban, valamint abban, hogy az  $n$ -gramokat milyen arányban találja meg a program. Azt az eredményt kaptuk, hogy a szavak és szópárok találati arányában szinte alig van különbség, a szavak találati aránya mindkét fajta szövegben a 100%-hoz közelít, a szópároké pedig 80-90% között van. A 3-gramok és 4-gramok találati arányában van a legnagyobb következetes különbség. Az 5-gramokban is nagy különbség van, de nem következetesen az emberi fordítás vagy a jobb gépi fordítás javára.

A találati számokban a különbség nem rendszeres a szavak, szópárok esetében, a 3-gramok átlagos találati eredményei viszont következetesen nagyobbak a jobb szövegek esetében. A 4-gramok és 5-gramok is hasonló tendenciát mutatnak, de itt jóval több a kivétel.

Az eredmények alapján a módszert a következőképpen módosítottuk:

- A gyakoriságokat is figyelembe vesszük, azért, mert a gyakrabban előforduló  $n$ -gramok nagyobb elfogadhatóságot sugallnak, még ha nem is ez az elfogadhatóság fő kritériuma.
- A gyakoriságot nem lineárisan vesszük figyelembe, mert két  $n$ -gram találati számában a különbség nem feltétlenül jelent ugyanakkora különbséget az elfogadhatóságában.
- A súlyokat az előző eredmények alapján módosítjuk, méghozzá úgy, hogy 3 hosszúságú  $n$ -gramok súlya legyen a legnagyobb.
- A súlyok módosítása után a jó szövegek is alacsonyabb értéket fognak kapni, így az eredményt végül egy szigmoid jellegű függvénnyel az intervallum teljes szélességében „kiterjesztjük”: a jó szövegek jobb, a rossz szövegek pedig rosszabb értéket kapjanak

A módosítások után az eredmények a következők:

Szöveg	Eredmény
Rossz szövegek:	0,094
Gépi fordítás:	0,335
Emberi szövegek:	0,849

9. táblázat: A tesztszövegek átlagos minősítési eredménye a végső módszerrel

#### 5.4.3. Emberi értékelés – gépi értékelés

Minden automatikus értékelési módszer jóságának végső ismérve az, hogy az általa adott eredmények mennyire korrelálnak az emberi ítélettel. Így tehát az általunk létrehozott módszert is össze kell hasonlítani az emberi értékeléssel. A módszer érvényességének vizsgálatára a kísérleti értékelés után kérdőíves felmérés készült, hogy a gépi értékelést és az emberi értékelést össze lehessen hasonlítani.

A felmérésben az alanyoknak értékelniük kellett a géppel fordított szövegeket, de nem skálán, mivel ez a laikus olvasók számára túlságosan nagy nehézséget jelentene, valamint az eredmények megbízhatósága sem garantált. A Callison-Burch által lefolytatott mérés tapasztalataira is támaszkodva, ahol a számskálás értékelés helyett a korábbi tapasztalatok miatt a sorrendi értékelést tartották célravezetőbbnek, az alanyoknak ebben a vizsgálatban is sorrendet kellett felállítaniuk (Callison-Burch et. al. 2008: 75). A szövegeket nem tagoltuk részekre, hanem a szövegek egészét tekintettük egy egységnek.

A kérdőívhez kiválasztott szövegek szintén internetes újságcikkek gépi fordításai. A szövegek témája különböző, egy kivételével mindegyik általános, nem tekinthető szakszövegnek. A kérdőíven 7 géppel fordított szöveg szerepel, mindegyiknek a hossza nagyjából 200 szó. Ezeket kellett a jelölteknek érthetőség és olvashatóság szerint sorrendbe tenni. A szövegek között voltak szembetűnően különbözőek és csak árnyalatnyi különbségeket mutatók. A kérdőívet 41 fő töltötte ki. Ugyanezeket a szövegeket a végső, módosított módszerrel automatikusan is értékeltük. A szövegekre kapott eredmények a következők lettek:

Szöveg sorszáma	Átlagos emberi pontszám	Pozíció	Gépi pozíció
1.	4,18	4.	4.
2.	6,10	7.	7.
3.	2,80	2.	2.
4.	2,65	1.	3.
5.	4,25	6.	6.
6.	4,20	5.	5.
7.	3,83	3.	1.

**10. táblázat: Emberi és gépi értékelés sorrendje**

A táblázat megmutatja az emberek és a gép által felállított sorrendet. Az emberi és a gépi értékelés korrelációjának meghatározásához a Spearman korrelációs együtthatót használjuk (Spearman 1904: 77). Az együttható megmutatja, hogy a két változó, ebben az esetben a gépi értékelés eredménye és az emberi értékelés eredménye, hogyan befolyásolja egymást. Ha az együttható értéke magas, azaz 1-hez közelít, az annyit jelent, hogy a két változó értékének változása követi egymást. Tehát ha a géppel adott érték növekszik, akkor az emberi érték is növekszik. Az együttható értéke ebben az esetben  $\sigma = 0,96$ , nagyon magas, a két értékelési módszer nagyban korrelál.

## 5.5. A METEOR értékelési módszer

A módszer alapötletét a korábban említett, referencia-fordításon alapuló metrikus minősítési módszerek adták. Ezek közül a METEOR módszer figyelembe veszi az illeszkedő összefüggő szakaszok hosszát is (a BLEU-vel ellentétben), így ígéretes alpnak tűnt az értékelési módszer kidolgozásához. A minősítés algoritmusát és a számítási módszert azonban a METEOR módszerétől jelentősen eltérően határoztuk meg, méghozzá azért, mert a rendelkezésre álló referencia jellegében nagyon nagy az eltérés: egyrészt esetünkben a referencia nem a jelölt szöveg fordítása, másrészt pedig a referencia nem egy konkrét szöveg, hanem egy korpusz.

Mindazonáltal megvizsgáltuk a METEOR módszert, és megpróbáltuk alkalmazni (alternatívaként) a bemutatott, általunk kidolgozott módszerrel ötvözve. A gépi fordítások értékelésében fontos szerepe van két mérőszámnak: a pontosságnak (precision) és a fedésnek (recall). A pontosság azt mutatja meg, hogy a jelölt fordításnak hány százaléka volt jó – azaz szerepelt a referenciafordításban is.

$$P = \frac{m}{w_i}$$

ahol  $m$  az egyező szavak száma, a  $w_i$  pedig a jelölt fordításban szereplő szavak száma. A fedés pedig azt adja meg, hogy a referenciafordítás hány százalékát sikerült „eltalálni” (Melamed et al. 2003: 1).

$$R = \frac{m}{w_r}$$

ahol  $m$  az egyező szavak száma, a  $w_r$  pedig a referenciafordításban szereplő szavak száma.

A METEOR módszer kifejlesztése során az eredmények azt bizonyították, hogy a fedés nagyobb jelentőségű, mint a pontosság, ha az emberi ítélettel korreláló mérőszámot szeretnénk kapni (Banerjee és Lavie 2005: 66). A BLEU indexnek pontosan az a hiányossága, hogy a fedést egyáltalán nem veszi figyelembe, helyette az  $n$ -gramok hosszát veszik figyelembe. A METEOR a BLEU-nek azt a hiányosságát is pótolja, hogy a nem a helyes szórendet jutalmazza, méghozzá úgy, hogy az egyező, összefüggő sorozatot alkotó szavakat jutalmazza hossz alapján (Banerjee és Lavie 2005: 67). Ehhez egy töredezettségi mérőszámot alkalmaznak, amellyel a rövidebb sorozatokat büntetik (Denkowski és Lavie 2010: 106). Miután az algoritmus megtalálta a szegmensben levő illeszkedő unigramokat, ezekből a lehető leghosszabb, egymást követő unigramokból álló sorozatokat készíti, és megadja ezek számát. Az általános képlet az egy szegmensre vonatkozó büntetésre a következő:



$$B = \gamma \cdot \left( \frac{s}{m} \right)^\beta$$

ahol  $s$  a szegmensben levő leghosszabb összefüggő sorozatok száma,  $m$  az illeszkedő unigramok száma,  $\beta$  és  $\gamma$  pedig szabad paraméterek, amelyek változtathatóak annak érdekében, hogy a mérőszámot az emberi ítélethez hangoljuk. A METEOR rendszerben empirikus meghatározás után  $\beta = 3$ ,  $\gamma = 0,5$  volt (Banerjee és Lavie 2005: 68), amelyet későbbi hangolás után az angol nyelvre  $\beta = 0,83$   $\gamma = 0,28$  számokban határoztak meg (Denkowski és Lavie 2010: 109). A szegmens végső minőségét a pontosság és a fedés súlyozott harmonikus átlagából és a büntetésből kapjuk meg. A harmonikus átlag:

$$F_{mean} = \frac{P \cdot R}{\alpha \cdot P + (1 - \alpha) \cdot R}$$

ahol  $P$  a pontosság,  $R$  a fedés,  $\alpha$  pedig az algoritmus harmadik szabad paramétere. Ennek értéke kezdetben  $\alpha = 0,9$ , később pedig  $\alpha = 0,81$  (Denkowski és Lavie 2009: 110,112).

A METEOR érték:

$$M = F_{mean} \cdot (1 - B)$$

ahol a  $B$  a korábban említett büntetés.

Ha már eleve létezik több ilyen minősítő módszer is, akkor felmerül a kérdés, hogy miért van szükség egy újabbra? Az új értékelési módszerre több okból is szükség van. Egyrészt a jelenlegi minősítés célja sem az, mint a korábban említett módszereké. A BLEU, NIST, METEOR minősítésének célja a szövegek általános minőségének mérése, legfőképp a pontosság és a gördülékenysége (Banerjee 2009: 4). Az általunk kidolgozott módszer az emberi szövegekhez való általános hasonlóságot, ezen keresztül pedig az elfogadhatóságot méri. Az emberi szöveghez való hasonlóságot hozzájárul a gördülékenységhoz is, tehát az módszer a két fő szempont közül a gördülékenységet méri. A pontosság mérésére teljesen alkalmatlan, valamint a korábban említett és szintén nagyon fontos tényezőt, az

érthetősége sem vizsgálja. Ez egyrészt a módszer hátránya is, másrészt pedig így lehetővé teszi, hogy a korábbi, érthetőségi mérésekkel összehasonlítva megállapítsuk, hogy mi az összefüggés az érthetőség és az elfogadhatóság között.

Másrészt a probléma ezekkel a minősítő módszerekkel az, hogy egy szöveget csak akkor tudnak értékelni, ha van hozzájuk legalább egy, de lehetőleg több emberi referenciafordítás. Mi olyan módszert szerettünk volna létrehozni az elfogadhatóság vizsgálatára, amelyhez nincsen szükség emberi fordításra, és amellyel gyorsan minősíthető bármilyen szöveg. A minősítés legfontosabb célja pedig az, hogy mérni tudjuk a szövegek javulásának mértékét. A fent említett módszerek célja inkább a fordítórendszerek értékelése, míg az általunk létrehozott módszer a szövegekre koncentrál. Erre a célra kézenfekvő az internet szövegeinek használata a referenciafordítások helyett. A referenciafordítások jellegéből adódóan viszont a fent említett számítási módszerek a mi mérési algoritmusunkban nem alkalmazhatóak.

A saját módszerünk alkalmazható bármilyen géppel fordított szöveg minősítésére, vagy akár nyelvtanulók szövegeinek minősítésére, de csak az elfogadhatóság szempontjából. Egyelőre csak magyar nyelvű szövegekre teszteltük, és magyar nyelvre van hangolva, de feltételezésünk szerint más nyelvekre is működik kis változtatásokkal. Arra feltétlenül vigyázni kell, hogy az olyan nyelvek esetében, mint az angol, amelyet sok nem anyanyelvi beszélő használ, az interneten megtalálható szövegeknek esetleg nagy része nem elfogadható. Ezeket az aspektusokat és a magyaron kívül más nyelveket nem vizsgáltunk.

## **5.6. A METEOR számítási módszer problémája**

### **5.6.1. A töredezettségi büntetés**

A METEOR minősítési módszer részletes leírására azért volt szükség, hogy be tudjuk mutatni a módszernek azt a részét, amellyel nem értünk teljes mértékben

egyét. Fent láthattuk, hogy a módszer büntetést rendel a szegmensekhez aszerint, hogy a jelölt és a referencia fordításban illeszkedő részei mennyire töredeztettek (vagy mennyire összefüggőek). A módszer elve az, hogy az illeszkedő unigramsorozatok minél több összefüggő részből állnak, azaz az összefüggő részek minél rövidebbek, a szegmens annál nagyobb büntetést érdemel. A büntetés mértékét úgy számítják ki, hogy az összefüggő szakaszok számát elosztják a szegmensben levő összes illeszkedő unigramszámmal:  $B = \gamma \cdot \left(\frac{s}{m}\right)^\beta$

ahol B a töredezettségi büntetés, s a szegmensben levő leghosszabb összefüggő sorozatok száma, m az illeszkedő unigramok száma,  $\beta = 3$  és  $\gamma = 0,5$ . A teljes minősítési képlet figyelembe veszi ugyan a fedést és a pontosságot, és a büntetés nagyságrendje a harmadik hatvány miatt igen kicsi, de úgy vettük észre, hogy a büntetés nem egészen a jóság alapján rendezi sorba a szegmenseket. Nézzünk egy példát:

Egy 10 unigramból (továbbiakban szóból) álló szegmens néhány lehetséges felosztása összefüggő szakaszokra a következő:

Szakaszok				Fedettség	Szakaszok száma	METEOR büntetés
10				10	1	0,100
9				9	1	0,111
8				8	1	0,125
7				7	1	0,143
6				6	1	0,167
5				5	1	0,200
1	9			10	2	0,200
2	8			10	2	0,200
3	7			10	2	0,200
4	6			10	2	0,200
5	5			10	2	0,200
1	8			9	2	0,222
3	6			9	2	0,222
4	5			9	2	0,222
4				4	1	0,250
1	7			8	2	0,250

Szakaszok				Fedettség	Szakaszok száma	METEOR büntetés
2	6			8	2	0,250
3	5			8	2	0,250
4	4			8	2	0,250
1	6			7	2	0,286
2	5			7	2	0,286
3	4			7	2	0,286
1	1	8		10	3	0,300
1	3	6		10	3	0,300
1	4	5		10	3	0,300
2	2	6		10	3	0,300
2	3	5		10	3	0,300
2	4	4		10	3	0,300
3	3	4		10	3	0,300
3				3	1	0,333
1	5			6	2	0,333
2	4			6	2	0,333
3	3			6	2	0,333
1	1	7		9	3	0,333
1	2	6		9	3	0,333
1	3	5		9	3	0,333
1	4	4		9	3	0,333
2	2	5		9	3	0,333
2	3	4		9	3	0,333
3	3	3		9	3	0,333
1	1	6		8	3	0,375
1	2	5		8	3	0,375
1	3	4		8	3	0,375
2	2	4		8	3	0,375
2	3	3		8	3	0,375
1	4			5	2	0,400
2	3			5	2	0,400
1	1	1	7	10	4	0,400

**11. táblázat: 10 szóból álló szegmens néhány felosztása a büntetés értékével.**

(A 11.4 mellékletben megtalálható a 10 szóból álló szegmens minden felosztása a METEOR büntetés és a legvégső javított büntetés értékével.)

A táblázat első 10 oszlopa jelzik az illeszkedő, összefüggő szakaszok hosszát, a 11. oszlop az illeszkedő szavak száma, a 12. oszlop pedig a büntetés mértéke. A

táblázatból láthatjuk, hogy a büntetés mértéke egyenlő akkor, ha az illeszkedő szavak nem alkotnak összefüggő szakaszokat, tehát csak a szavak illeszkednek nem a megfelelő sorrendben. A büntetés mértéke nincs tekintettel arra, hogy az illeszkedés mértéke mekkora. Tehát ha a 10 szóból összesen 2 illeszkedik, vagy ha a 10 szóból 9 illeszkedik, de az eredetitől eltérő sorrendben – tehát minden szakasz 1 hosszúságú – , akkor a büntetés mértéke egyenlő, pedig láthatjuk, hogy a kettő között hatalmas különbség van. Ez főleg igaz a magyarra, ahol a szórend nem kötött, így az illeszkedő szavak sorrendje nem annyira fontos. Ezért az 2/2-es illeszkedés nagyobb büntetést érdemelne, mint a 9/9-es.

A másik, azonnal látható probléma az, hogy a jobban illeszkedő, de rövidebb szakaszokat is tartalmazó felbontások nagyobb büntetést kapnak, mint a kevésbé illeszkedőek. A 4 + 5 hosszúságú szakaszokból álló felbontás nagyobb büntetést kap, mint a 4+5+1 szakaszokból álló, pedig ugyanaz a két, hosszú illeszkedő szakasz mindkettőben szerepel, csak a másodikban ráadásul a maradék 1 szó is illeszkedik, tehát egyértelműen kisebb büntetést érdemelne. Ez a probléma más esetekben is fennáll, még ennél szélsőségesebb esetekre is. Fedettségnek hívjuk az illeszkedő szavak számát, felosztásnak pedig az összefüggő szakaszokat. A következő táblázat egy 20 szóból álló szegmens néhány büntetését mutatja:

Felosztás						Fedés	Nem fedett	Szakaszok száma	Szakaszok/ fedettség	Meteor büntetés
4	4					8	12	2	0,25	0,007813
4						4	16	1	0,25	0,007813
1	3	5	5	5		19	1	5	0,263158	0,009112
1	4	4	5	5		19	1	5	0,263158	0,009112
2	2	5	5	5		19	1	5	0,263158	0,009112

**12. táblázat: 20 szóból álló szegmens néhány felosztása a METEOR büntetés értékével.**

Tehát a METEOR büntetése túlságosan jutalmazza a kevés szakaszból álló illeszkedést a lefedettséggel szemben. Ennek a problémának az orvoslására

megpróbáltunk kifejleszteni egy saját képletet, az eredeti büntetési képlet felhasználásával.

Első szempontként a METEOR büntetést úgy kell módosítani, hogy a fedettséget jobban vegye figyelembe. Ehhez hozzávettük szorzóként a szegmensben levő nem lefedett szavak számát. De így a büntetés túlzottan jutalmazza a lefedettséget, és nem veszi figyelembe, hogy az eredeti szegmens milyen hosszú, valamint hogy milyen hosszúak az összefüggő szakaszok. A lefedetlen részeknek növelniük kell a büntetést, a hosszak pedig csökkentenie, hiszen minél hosszabb a szegmens, és minél több ebben a lefedett rész, annál jobbnak kell, hogy számítson. A büntetésbe szintén bele kell, hogy számítsa, hogy az összefüggő szakaszok milyen hosszúak, tehát minél hosszabbak ezek a szakaszok, annál kisebb a büntetés.

Ezért a számot még arányosítottuk a szegmens hosszával, plusz a szegmensben levő leghosszabb szakasz hosszával. Így tehát a következő képletet hoztuk létre, amely az előzetes tesztek alapján megfelelően bünteti a töredezettséget, de figyelembe veszi a lefedettséget is. Az új arány tehát a következő:

$$A = \frac{s + n}{m + \max + h},$$

ahol  $s$  az összefüggő szakaszok száma,  $n$  a nem lefedett szavak száma,  $m$  a lefedett szavak szám,  $max$  a leghosszabb szakasz hossza,  $h$  pedig a szegmens teljes hossza. A büntetés meghatározásához a METEOR képletben szereplő szabad paraméterek értékeként a 2005-ös kutatás eredményeit használtuk, mivel egyrészt Denkowski és Lavie 2010-es felmérésében csak az angol nyelvre hangolták ezeket a paramétereket, másrészt pedig itt nem a paraméterek beállítása a cél, hanem a büntetés értékének olyan meghatározása, amely eleve megfelelően állítja sorrendbe a szegmens felosztásait. Így tehát a  $\beta = 3$  és  $\gamma = 0,5$  értéket használjuk. A büntetés tehát:

$$B = 0,5 \cdot \left( \frac{s + n}{m + max + h} \right)^3$$

ahol ugyanúgy  $s$  az összefüggő szakaszok száma,  $n$  a nem lefedett szavak száma,  $m$  a lefedett szavak szám,  $max$  a leghosszabb szakasz hossza,  $h$  pedig a szegmens teljes hossza. Az új mérték megfelelő olyan szempontból is, hogy a nagyságrendje a METEOR büntetéssel nagyrészt megegyezik, így egyszerűen használható az eredeti büntetési érték helyett. A következő táblázat mutatja a felosztások közül az előzőleg kiemelt sorokra a büntetés értékét először 10 szóból, majd 20 szóból álló szegmens esetében.

Meteor sorrend	Új sorrend				Fe-dés	Nem fedett	Szakaszok száma	METEOR büntetés	Új büntetés
25	12	1	4	5	10	0	3	0,01350	0,000864
14	16	4	5		9	1	2	0,00549	0,000977
48	18	1	1	1	7	10	4	0,03200	0,001626
34	20	1	1	7		9	3	0,01852	0,001821
4	58	7			7	3	1	0,00146	0,002315
22	60	3	4		7	3	2	0,01166	0,006749
33	86	3	3		6	4	2	0,01852	0,015746
47	105	2	3		5	5	2	0,03200	0,029407
15	120	4			4	6	1	0,00781	0,029407
30	130	3			3	7	1	0,01852	0,062500

13. táblázat: A felosztások sorrendje és a büntetés az új képlettel 10 szó esetében.

A táblázatból látható, hogy az új büntetési mértékkel a felosztások megfelelő sorrendben vannak, ha a jóság szempontja az, ha a szakaszok minél hosszabbak, és a lefedettség minél nagyobb. A sorrendbeli változás egyes esetekben egészen nagy, például az egy darab, 3 szóból álló szakaszt tartalmazó szegmens az eredeti 30. pozícióból a 130. pozícióba ugrott, ami az összes felbontást tekintve a megfelelő pozíció, hiszen ez elég kis fedést jelent a 10 szavas szegmensben. Az 1 + 4 + 5 hosszú szakaszokat tartalmazó, teljesen fedett szegmens a 25. helyről a 12. helyre került, ami szintén jó, hiszen a szakaszok száma kicsi, a szakaszok hosszúak, és a teljes szegmens lefedik.

Meteor sorrend	Új sorrend						Fe-dés	Nem fedett	Szakaszok száma	METEOR büntetés	Új büntetés
32	21	1	3	5	5	5	19	1	5	0,009112	0,001268
33	22	1	4	4	5	5	19	1	5	0,009112	0,001268
34	25	2	2	5	5	5	19	1	5	0,009112	0,001268
30	1030	4	4				8	12	2	0,007813	0,04187
31	1168	4					4	16	1	0,007813	0,111903

**14. táblázat: A felosztások sorrendje és a büntetés az új képlettel 20 szó esetében.**

Már ebből a néhány példából is jól látszik, hogy a METEOR büntetési képletével szinte egymás mellett álló felbontások milyen távol kerültek egymástól, a megfelelő helyre. A cél pontosan ez volt, a kevésbé lefedett, de viszonylag hosszú szakasz(ok)ból álló szegmenseket szerettük volna hátrább sorolni a jobban lefedettekhez képest. Hiszen könnyen belátható, hogy a 3 db 5 hosszúságú + 2 db 2 hosszúságú szakaszokból álló, szinte teljesen fedő felbontás sokkal jobb, mint az összesen egy darab 4 hosszúságú szakaszt tartalmazó felbontás. Az új képlet pontosan ezt az eredményt hozza, míg a METEOR még kisebb büntetést is adott az utóbbinak.

Sok olyan felbontás van, amelyek között szubjektíven is nehéz különbséget tenni, azaz nehéz megállapítani, melyik a jobb, ha valamelyikről elmondható



egyáltalán, hogy jobb. Ezt láthatjuk a fenti táblázat első három sorában is. Ezek a felbontások az új büntetési képlettel is ugyanakkora értéket kapnak.

A felbontások teljes táblázatában a büntetési értékeket megfigyelve észrevehetjük, hogy viszonylag sok felbontás kap így egyenlő mértékű büntetést. Ezek között a felbontások között nehéz különbséget tenni, mert amelyiknél nagyobb a fedettség, ott nagyobb a töredezettség is, és fordítva. Ha különbséget szeretnénk tenni köztük, akkor a fedettség és a kisebb töredezettség közül valamelyiknek nagyobb prioritást kell adni. A METEOR módszer eleve a kevésbé töredezett szegmenseknek kedvez a jobban fedettekkel szemben. Ha a jelölt és a referencia szegmensnek hosszabb részei egyeznek, ez valóban azt jelenti, hogy a jelölt szegmens jobban hasonlít az emberi fordításhoz, a szavak sorrendje a szokásos, tehát elfogadhatóbb is. De ha ezzel együtt a szegmens többi része egyáltalán nem is illeszkedik, tehát még szinonimák szintjén sem, akkor emiatt a szegmens eltér az elfogadhatótól. (A METEOR módszer illeszkedő szinonimákat is keres (Banjeree és Lavie 2005: 67).) Ha a két mérték (a szakaszok hossza és a fedettség) kétféle felbontásnál nagyjából kiegyenlíti egymást, akkor nem jelent problémát, ha a büntetésük mértéke megegyezik. A kérdés az, hogy hol van az a határ, ahol még kiegyenlítik egymást, és valóban egyenlő büntetést érdemelnek. Nézzünk egy példát: egy 20 szó hosszúságú szegmens esetében az (a) 1+1+1+5+5 felbontás és a (b) 2+2+2+2+2+2+2 felbontás egyforma büntetést kap. Az (a) felbontásnál a fedettség csak 13, míg a (b) esetében 16, viszont láthatjuk, hogy a (b) felbontásban több rövidebb szakasz van, tehát nagyobb a töredezettsége. A kérdés az, hogy vajon ezek tényleg egyforma büntetést érdemelnek? Ha nem, akkor melyik felbontás a rosszabb?

A következő táblázat néhány azonos értékű felbontást mutat:

Felbontás									Fedettség	Büntetés
1	1	1	5	5					13	0,015746
1	1	2	4	5					13	0,015746

1	1	3	3	5					13	0,015746
1	2	2	3	5					13	0,015746
2	2	2	2	5					13	0,015746
1	1	1	3	4	4				14	0,015746
1	1	2	2	4	4				14	0,015746
1	1	2	3	3	4				14	0,015746
1	2	2	2	3	4				14	0,015746
2	2	2	2	2	4				14	0,015746
1	1	1	3	3	3	3			15	0,015746
1	1	2	2	3	3	3			15	0,015746
1	2	2	2	2	3	3			15	0,015746
1	2	3	3	2	2	2			15	0,015746
2	2	2	2	2	2	3			15	0,015746
2	2	2	3	2	2	2			15	0,015746
2	2	2	2	2	2	2	2		16	0,015746

**15. táblázat: Azonos értékű felbontások 20 szóból álló szegmensben**

Ha a fedettségben csak kicsi, például 1 szónyi különbség van, a töredezettségben pedig jelentős a különbség, akkor a büntetés megfelelő lesz, ha a kevésbé fedett, de kevésbé töredezett felbontásra kisebb az értéke, vagy akár egyenlő a két felbontásra. Véleményünk szerint ha viszont már ilyen nagy a különbség a fedettségben, mint (a) és (b) példa esetében (4 szó), ekkor már különbséget kell tennünk a két felbontás között, méghozzá a jobban fedett javára.

Ugyanez a probléma fennáll arra az esetre is, ahol a töredezettség nagyon nagy, de a fedettség is nagyon nagy. Az ilyen felbontásokat ez a képlet is még mindig túlságosan bünteti a kevésbé töredezett, de jóval kisebb fedettséget adó felbontásokkal szemben.

															Fedettség	Büntetés
2	5	5													12	0,0131
3	4	5													12	0,0131
1	4	4	4												13	0,0131
2	3	4	4												13	0,0131
3	3	3	4												13	0,0131
2	3	3	3	3											14	0,0131
1	1	1	1	1	5	5									15	0,0135
1	1	1	1	2	4	5									15	0,0135
1	1	1	1	3	3	5									15	0,0135

1	1	1	2	2	3	5							15	0,0135
1	1	2	2	2	2	5							15	0,0135
1	1	1	1	1	3	4	4						16	0,0135
1	1	1	1	2	2	4	4						16	0,0135
1	1	1	1	2	3	3	4						16	0,0135
1	1	1	2	2	2	3	4						16	0,0135
1	1	2	2	2	2	2	4						16	0,0135
1	1	1	1	1	3	3	3	3					17	0,0135
1	1	1	1	2	2	3	3	3					17	0,0135
1	1	1	2	2	2	2	3	3					17	0,0135
1	1	1	2	3	3	2	2	2					17	0,0135
1	1	2	2	2	2	2	2	3					17	0,0135
1	1	2	2	2	3	2	2	2					17	0,0135
1	1	2	2	2	2	2	2	2	2				18	0,0135
1	1	1	1	1	1	1	1	1	4	5			18	0,0138
1	1	1	1	1	1	1	1	2	3	5			18	0,0138
1	1	1	1	1	1	1	2	2	5				18	0,0138
1	1	1	1	1	1	1	1	1	2	4	4		19	0,0138
1	1	1	1	1	1	1	1	1	3	3	4		19	0,0138
1	1	1	1	1	1	1	1	2	2	3	4		19	0,0138
1	1	1	1	1	1	1	2	2	2	2	4		19	0,0138
1	1	1	1	1	1	1	1	1	2	3	3	3	20	0,0138
1	1	1	1	1	1	1	1	2	2	2	3	3	20	0,0138
1	1	1	1	1	1	1	1	3	3	2	2	2	20	0,0138
1	1	1	1	1	1	1	2	2	2	2	2	3	20	0,0138
1	1	1	1	1	1	1	2	2	3	2	2	2	20	0,0138

**16. táblázat: Kis fedettség, kis töredezettség – nagy fedettség, nagy töredezettség**

Az adatokból látható, hogy például a 12-es fedettségű 2+5+5 felbontás büntetése nagyobb, mint a 20-as fedettségű 1+1+1+1+1+1+2+2+3+2+2+2 felbontásé. Erre a jelenségre sok példát láthatunk, ha megnézzük az összes felbontást. Ez problémát jelent, tehát mindenképpen módosítani kell a képletet. Tehát a büntetés kiszámításakor a fedettségnek nagyobb súlyt kell adnunk, és így módosíthatjuk a büntetés mértékét. Ezenkívül a fedettségnek még nagyobb prioritást adhatunk anélkül, hogy a büntetés nagyságrendjén változtatnánk, méghozzá úgy, hogy a nevezőben levő  $n-t$  (a nem fedett szavak számát) szintén súlyozzuk. Így a fenti két problémára megoldást tudunk adni. Egyrészt a felbontások büntetése nagyobb

változatosságot mutat, másrészt pedig a jobb fedettségű felbontások kisebb büntetést kapnak.

$$B = 0,5 \cdot \left( \frac{s + \beta \cdot n}{\alpha \cdot m + \max + \beta \cdot h} \right)^3$$

A súlyok változtatható paraméterek, az előzetes teszteknek megfelelően jelenleg  $\alpha = 2$ ,  $\beta = 2$ . Ezzel az új, módosított képlettel a büntetések mértéke az előző táblázatban szereplő felbontásokra a következő:

															Fedettség	Büntetés
1	1	1	1	1	1	1	1	1	2	3	3	3			20	0,00192
1	1	1	1	1	1	1	1	2	2	2	3	3			20	0,00192
1	1	1	1	1	1	1	1	3	3	2	2	2			20	0,00192
1	1	1	1	1	1	1	2	2	2	2	2	3			20	0,00192
1	1	1	1	1	1	1	2	2	3	2	2	2			20	0,00192
1	1	1	1	1	1	1	1	1	2	4	4				19	0,00249
1	1	1	1	1	1	1	1	1	3	3	4				19	0,00249
1	1	1	1	1	1	1	1	2	2	3	4				19	0,00249
1	1	1	1	1	1	1	2	2	2	2	4				19	0,00249
1	1	2	2	2	2	2	2	2							18	0,00289
1	1	1	1	1	1	1	1	1	4	5					18	0,00318
1	1	1	1	1	1	1	1	2	3	5					18	0,00318
1	1	1	1	1	1	1	2	2	2	5					18	0,00318
1	1	1	1	1	3	3	3	3							17	0,00370
1	1	1	1	2	2	3	3	3							17	0,00370
1	1	1	2	2	2	2	3	3							17	0,00370
1	1	1	2	3	3	2	2	2							17	0,00370
1	1	2	2	2	2	2	2	3							17	0,00370
1	1	2	2	2	3	2	2	2							17	0,00370
1	1	1	1	1	3	4	4								16	0,00467
1	1	1	1	2	2	4	4								16	0,00467
1	1	1	1	2	3	3	4								16	0,00467
1	1	1	2	2	2	3	4								16	0,00467
1	1	2	2	2	2	2	4								16	0,00467
1	1	1	1	1	5	5									15	0,00582
1	1	1	1	2	4	5									15	0,00582
1	1	1	1	3	3	5									15	0,00582
1	1	1	2	2	3	5									15	0,00582
1	1	2	2	2	2	5									15	0,00582
2	3	3	3	3											14	0,00686

1	4	4	4										13	0,00850
2	3	4	4										13	0,00850
3	3	3	4										13	0,00850
2	5	5											12	0,01044
3	4	5											12	0,01044

**17. táblázat: Ugyanazon felbontások a módosított büntetéssel**

A táblázat az új büntetési értékek alapján van sorba rendezve, és láthatjuk, hogy így a kisebb fedettségű felbontások nagyobb büntetést kapnak, tehát a sorrend megfelelő lesz. A büntetést meghatározó képletet tehát megfelelőnek fogadhatjuk el.

## 5.6.2. A büntetés alakítása az internetes értékelési módszerhez

### 5.6.2.1. A módosítás szükségessége

A képlet létrehozása és a felbontások „elméletbeni” tesztelése után célszerűnek látszik a módszer kipróbálása és alkalmazása az általunk létrehozott internetes értékelési módszerhez. Ehhez természetesen az értékelés drasztikus módosítására van szükség, hiszen a súlyozás helyett a büntetési rendszerrel kell meghatározni az egyes szegmensek, majd ezek után a szöveg minőségét.

Az általunk kidolgozott értékelési módszer alapja az, hogy az értékeléshez használandó referenciaszövegekként az interneten található magyar szövegeket használjuk, és ezt az alapot továbbra is megtartjuk, hiszen ez adja a módszer egyediségét. Ezzel biztosítjuk, hogy továbbra sem lesz szükség költségesen, kézzel létrehozott referenciaszövegekre. A METEOR módszer viszont referenciaszövegekként feltételezi a jelölt szöveg fordítását. Így könnyen kiszámolható a fedés és a pontosság is, valamint viszonylag egyszerűen összeállíthatók az összefüggő, fedő szakaszok. Ezek viszont problémát jelentenek, ha a referencia egy korpusz.

### 5.6.2.2. A fedés problémája

Ha a referencia viszont nem a szöveg fordítása, ráadásul nem is adott hosszúságú, a fedés kiszámítása nehézségeket jelent, ha egyáltalán értelmezhető ez a fogalom. Hiszen a fedés azt jelenti, hogy a referenciaszövegnek hány százalékát sikerült „eltalálni” az értékelendő fordításban. Tehát a METEOR módszer alkalmazása a „végtelen” hosszúságú – de legalábbis meghatározhatatlan hosszúságú – referenciaszövegekre nehézségeket okoz.

Fedés meghatározáshoz tehát egy teljesen más, nem szokványos megközelítést alkalmaztunk: azokat a szavakat vesszük bele a fedésbe, amelyek szerepelnek egy olyan, legalább 3 hosszúságú  $n$ -gramban, amelyre volt találat. Az egyszerűség kedvéért hívjuk ezeket megtalált szavaknak, és ezeknek a számát jelöljük  $f$ -fel. A szegmensben található összes elem a szegmens hossza, a fentiek alapján jelöljük  $h$ -val. A fedés tehát:

$$R = \frac{f}{h}$$

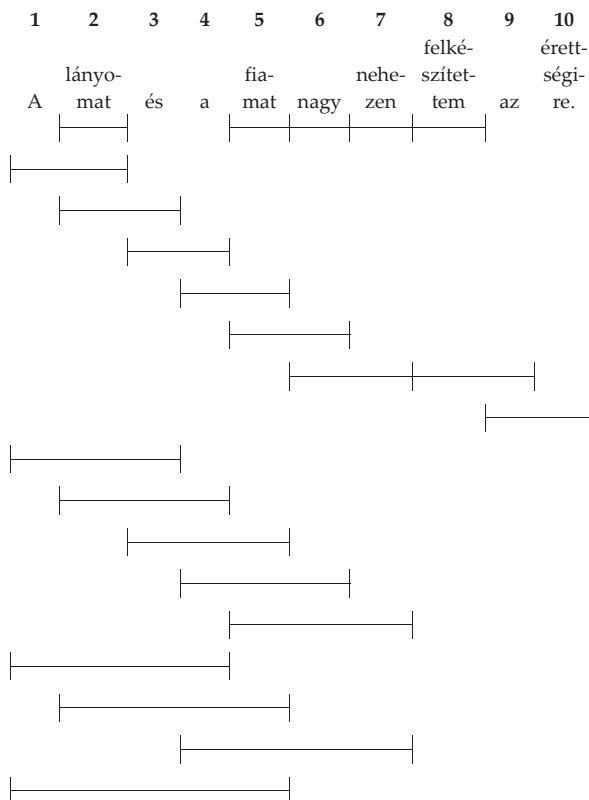
A pontosság kiszámításához pedig a szegmensben megtalált szavak számát, tehát a fedettséget használjuk. A korábbi jelölésnek megfelelően  $m$  a megtalált szavak száma,  $h$  pedig a szegmens hossza.

$$P = \frac{m}{h}$$

### 5.6.2.3. Szakaszok kialakítása

Szintén nehézséget okoz annak meghatározása, hogy az összefüggő szakaszokat hogyan állítsuk össze. Ha csak egyetlen referenciaszöveg áll a rendelkezésünkre, a lefedettségéből átfedés nélkül kialakíthatóak az összefüggő szakaszok. De a mi helyzetünkben a megtalált összefüggő szakaszok átfedésben lehetnek, sőt, nagyon gyakran átfedésben vannak. Ezek közül tehát meg kell találni a legjobb átfedést. Nézzünk egy példát. A példamondat a következő: (8) *A lányomat és a fiamat nagy nehezen felkészítettem az érettségire.* A példamondat keresési eredményeit a

következő ábrán láthatjuk. Minden szó egy egységet jelent az egyenesen, a szakaszok pedig a megtalált n-gramok. (Az 1 és 2 hosszú szavakat nem keressük.)



**3. ábra: A megtalált szakaszok a példamondatban.**

Az ábrán látszik, hogy a szegmens teljesen lefedett, tehát minden szóra van találat. Az is egyértelműen látszik, hogy egyetlen 5 hosszúságú szakasz van, az első 5 szóra. Ez az 5 hosszúságú szakasz tehát használható is a szegmens első felére. A második felét viszont csak egymással átfedő szakaszokból tudjuk összeállítani.

A legjobb lefedés összeállításakor két szempontot kell szem előtt tartani: egyrészt a lehető leghosszabb szakaszokat kell kiválasztani, másrészt pedig a lehető legtöbb szót le kell fedni. Mindkét szempont egyformán fontos, de egymásnak részben ellentmondanak. Ha kiválasztjuk az egymással nem átfedő szakaszokat, akkor ezek nem feltétlenül a leghosszabbak, és nagy valószínűséggel nem is biztosítják a teljes lefedettséget. A leghosszabb szakaszokat is kiválaszthatjuk, de ezek átfedhetnek egymással, vagy az is előfordulhat, hogy nem fedik le a szegmenst (még akkor sem, ha egyébként minden szó illeszkedik). Mindkét esetben az átfedő szakaszokat is fel kell használni, a kérdés az, hogy az átfedést hogyan kezeljük.

Az egyik lehetőség a lefedés megállapítására a leghosszabb és legtöbbet lefedő szakaszok kiválasztása úgy, hogy minden szó, amelyre volt találat, le legyen fedve. Így azonban a szakaszok hosszának összege nagyobb lesz, mint a teljes szegmens hossza, amit lehetne teljes fedésként kezelni, de a valódi fedésről mégis torz képet ad. A jelenlegi példában ez a felbontás a következő lenne:  $5 + 4 + 2 + 2 = 13$

A másik lehetőség az, hogy kiválasztjuk az összes szakasz közül a lehető leghosszabbakat, amelyek nem fednek át, és a lehető legnagyobb fedést biztosítják, majd megnézzük, hogy az eddig nem lefedett szakaszokra van-e illeszkedő szakasz, és ezt olyan hosszúnak vesszük, amekkora a kimaradó rész volt. Ez sem egészen a valóságot tükrözi, hiszen ez a módszer az előbb említettel szemben nagyobb töredezettséget fog adni a valódinál. Ez a felbontás a fenti példában:  $5 + 2 + 2 + 1$ .

Ez az utóbbi megoldás viszont jobban tükrözi a valóságot, valamint megadja a lehető legjobb fedettséget, tehát ezt fogjuk használni a programban is.

#### 5.6.2.4. Egy algoritmus a legjobb felbontás megkeresésére

A legjobb felbontás megkeresésének a problémáját a szegmensekben a következőképpen fogalmazhatjuk meg. A szegmenst képzeljük el egész számok



egy intervallumaként. Az intervallum kezdőpontja a szegmens eleje, azaz 1, az intervallum végpontja pedig a szegmensben szereplő szavak száma:  $h$ . A szegmensben megtalált  $n$ -gramok mindegyike egy-egy szakasz az intervallumon belül. Minden szakasz kezdőpontja a megtalált  $n$ -gram kezdete, a szakasz hossza pedig az  $n$ -gram hossza. Minden szakasz kezdőpontja nagyobb vagy egyenlő, mint az intervallum kezdőpontja, és minden szakasz végpontja kisebb vagy egyenlő, mint az intervallum végpontja. Ha az  $n$ -gramok hosszára nem adunk felső korlátot, akkor az intervallumban megtalálható szakaszok maximális száma  $q = h + (h-1) + (h-2) + \dots + 2 + 1 = \frac{h(h+1)}{2}$ . A szakaszok minimális hossza 1, maximális hossza  $h$ . A probléma megoldása azon szakaszok kiválasztása, amelyek egyrészt az intervallumnak a legjobb lefedését biztosítják, másrészt a lehető leghosszabb szakaszokat tartalmazzák, tehát amely a lehető legkisebb büntetési értéket adja a lehetséges felbontások közül.

A probléma tehát optimalizálási probléma. Optimalizálási problémákra megoldás lehet a dinamikus programozás, ha a probléma felbontható optimális részproblémákra. A mi esetünkben a problémát fel lehet bontani optimális részproblémákra, de a dinamikus programozás műveletigénye nagy, ezért megnézzük annak a lehetőségét is, hogy a feladat mohó algoritmussal megoldható-e. A mohó algoritmussal való megoldhatóságnak a feltétele, a mohó választási tulajdonság és az optimális részproblémákra bonthatóság (Cormen et al. 2001: 287-288). A második tulajdonsággal rendelkezik a probléma, de a mohó választási tulajdonságot meg kell vizsgálni.

Nem formálisan megfogalmazva optimális részproblémákra felbontható a feladat, mert ha egy szakasz részt vesz az optimális felbontásban, akkor a további feladat az lesz, hogy a tőle jobbra és balra levő intervallumokban is megkeressük az optimális felbontást.

Hasonló problémákra találhatunk már létező algoritmusokat, de ezeknél a problémáknál az optimális megoldás tulajdonságai mások, például a lehető

legkevesebb szakaszból álló lefedést kell megtalálni (a teljes lefedés eleve adott), vagy azt a lehető legjobb lefedést kell megtalálni, ahol a szakaszok nem átfedőek (pl. esemény-kiválasztási probléma, Cormen et al. 2001: 283). Nem sikerült olyan feladatot találni, ahol az optimális megoldás tulajdonságai ezzel megegyeznek, de a létező algoritmusok jó kiindulási pontot adnak.

A probléma többféle algoritmussal megoldható, de célunk az, hogy olyan algoritmust találjunk, amely polinomidőben megtalálja a legjobb megoldást. Nem polinomidejű megoldást könnyen lehet találni: minden lehetséges kombinációt ki kell értékelni, és a lehető legjobbbat választani, de mivel a hasonló jellegű problémákra vannak ennél jóval kisebb bonyolultságú megoldások, erre a feladatra is szeretnénk ilyet találni. Magától értetődő ötlet mindig a leghosszabb szakasznak a kiválasztása, de ennek bonyolultsága mindig  $O(q^2)$ , ráadásul az átfedések miatt minden választáskor változtatni kell az átfedésben részt vevő szakaszok hosszát. Ha a szakaszokat sorba rendezzük hossz alapján, akkor minden egyes választásnál módosítani kell az átfedésben érintett szakaszokat, utána pedig újra el kell végezni a rendezést. Ezek helyett inkább olyan algoritmussal próbálkozunk, amely a kezdőpont szerint sorba rendezett szakaszok sorozatában a választást az első szakasszal kezdi.

Az általunk kidolgozott algoritmus tetszőleges  $n$ -gram-hosszokkal (azaz tetszőlegesen hosszú szakaszokkal) is megtalálja a legjobb felbontást, valamint azt is kezelni tudja, ha a szakaszok végpontja az intervallum végpontjánál nagyobb, tehát a jelenlegi problémánál általánosabb problémákra is használható. A következőkben a szegmensre mint intervallumra, az  $n$ -gramokra mint szakaszokra, a szavakra mint az intervallum elemeire fogunk hivatkozni. Az intervallum elemei a szakaszok kezdőpontjai alapján rendezettek. Az algoritmus alapvetően a mohó stratégiát követi, de a megoldásban az egy lépéssel előbb megtalált és a megoldásba felvett szakaszt módosíthatja. Az algoritmus a következő:

Az első lépésben lecsökkentjük a szakaszok számát: az intervallum minden elemére csak az onnan kezdődő leghosszabbat tartjuk meg. Ezzel a lehetséges szakaszok számát maximum  $m$ -re csökkentük. A szakaszok számát tovább csökkenthetnénk úgy, hogy elvetünk minden olyan szakaszt, amelyet egy másik szakasz lefed. De mivel az algoritmus bonyolultságát ezek a szakaszok nem befolyásolják, ezt nem tesszük meg. (A jelenlegi feladatnál az adott kezdőponttal rendelkező leghosszabb szakaszok kiválasztása nem jelent plusz műveletet, mivel már az  $n$ -gramok keresésekor csak a leghosszabb találatokat tároljuk.)

A második feladat a maximum  $m$  darab szakaszból kiválasztani a legjobb lefedést biztosítóakat. Az intervallum minden eleménél megvizsgáljuk, hogy onnan milyen hosszú szakasz kezdődik, ez a hossz legyen  $h_i$ , a szakasz pedig  $m_i$ . Az  $i$ -edik szakasz kezdőpontja  $s_i$ , végpontja  $f_i$ . Ha  $i < j$ , akkor mindig igaz, hogy  $s_i < s_j$ . A felbontást alkotó szakaszokat  $l$ -lel indexeljük, és az aktuális szakaszt  $a_l$ -lel jelöljük. Ezután megvizsgáljuk, hogy a szakaszon belül kezdődik-e egy másik olyan szakasz, amely átlépi az  $i$ -edik szakasz végpontját, méghozzá legalább az aktuális  $m_i$  szakasz végpontja után kezdődő (az  $i + h_i$ -edik, azaz  $f_i + 1$ -edik pozíción) szakasz hossza  $+1$  hosszúságban:  $\exists h_j (j \in [s_i, f_i])$ , hogy  $f_j > f_i + h_i$ . Ha igen, ez azt jelenti, hogy ezt az új szakaszt ( $m_k$ ) érdekesebb lesz megtartani a megoldásban, mint az eredeti,  $m_i$  szakaszt és az utána következő,  $m_{i+h}$ -adikat. Ekkor összehasonlítjuk az  $m_i$  és az  $m_k$  szakasz hosszát, a rövidebb szakasz hosszából kivonjuk az átfedést: ha  $m_i \geq m_k$ , akkor  $h_k = h_k - (f_i - s_k)$ , különben  $h_i = h_i - (f_i - s_k)$ , és az  $h_i$ -t és az  $h_k$ -t felvesszük a fedettséget biztosító szakaszok közé: ezek lesznek  $a_i$  és  $a_{i+1}$ . Ezután ezt az új,  $m_k$  szakaszt ugyanígy megvizsgáljuk, azaz megnézzük, van-e benne olyan szakasz, amely átlépi a végpontot a fent említett feltételek szerint. Ha igen, akkor az előzőleg megjegyzett és felvett szakasz hosszát,  $a_{i+1}$ -et a fentiek szerint módosítjuk, ha hossza rövidebb, mint a benne levő, kellően hosszú szakasz hossza, és az új szakaszt szintén felvesszük a felbontásba. Ezt a folyamatot addig folytatjuk, amíg a szakaszban van olyan belső szakasz, amely átlép a végpontot.

Ha a szakaszban nem kezdődik a végpontot átlépő szakasz, akkor a szakaszt felvesszük a fedést biztosító szakaszok közé, és a szakasz végpontja utáni elemet kezdjük vizsgálni a fentiek szerint.

Ez az algoritmus lineárisan végigjárja egyszer az intervallum elemeit, és minden lépésben maximum egy visszalépést tartalmaz. A visszalépés nem valódi visszalépés, csak az előző lépés eredményét módosítja véglegesen, de a keresést a következő elemtől folytatja tovább. A szakaszokon belül kezdődő, a szakaszon túllépő szakaszok vizsgálata biztosítja a leghosszabb szakasz megtartását.

Ez az algoritmus egy eset kivételével mindig megtalálja a legjobb fedést biztosító szakaszokat. Az egy eset, amikor nem, az, amikor a szakaszban kezdődik olyan belső szakasz, amely saját magánál hosszabb, de nem átfedő része nem hosszabb, mint a szakasz után kezdődő szakasz hossza. Ekkor az algoritmus felveszi az eredeti szakaszt, és átugorja a szakaszban levő, esetlegesen jobb szakaszt. Ennek megoldására javasolunk egy másik, az intervallum végpontjától kezdő algoritmust, amely az előzőleg felvett megoldás módosítása nélkül, mohó stratégiával választja ki a megoldásban szereplő szakaszt. Ez a végponttól kezdő algoritmus a következő:

Az intervallum végpontjától indítjuk az elemek vizsgálatát az intervallumon visszafelé. Megkeressük azt a szakaszt, amely az adott elemet még lefedi, de az elemtől a lehető legtávolabb kezdődik, tehát az adott elemet lefedő leghosszabb szakaszt. Ezt a szakaszt felvesszük a lefedést biztosító szakaszok közé, a vizsgálatot pedig ennek a szakasznak a kezdőpontja előtti elemtől folytatjuk, amíg el nem érjük az intervallum elejét. Ez az algoritmus lényegesen egyszerűbbnek tűnik, műveletigénye mégis több, mint az első algoritmusé. Az intervallum elejétől induló keresés bonyolultsága  $O(m)$ , az intervallum végétől induló pedig maximum  $O(m^2)$ , mert azon kívül, hogy végigmegyünk az intervallumon, minden egyes elemre meg kell keresni a leghosszabb fedő szakaszt, amely az elemekre  $m + (m-1) + (m-2) + \dots + 1$  db vizsgálatot jelent. A mi esetünkben az  $n$ -gramok maximális hossza 5, így az algoritmus bonyolultsága  $O(m)$ -re csökken.

Ez az algoritmus szintén megtalálja a legjobb felbontást, szintén egy eset kivételével. A problémája ugyanaz, mint az első algoritmusnak, azaz ez az algoritmus is átugorja a szakaszon belüli esetleges jobb lehetőségeket.

A két algoritmus ötvözése viszont megoldást ad ezekre a hiányosságokra. Ha együtt alkalmazzuk őket, tehát először az előre haladó algoritmussal keressük a felbontást, és ha olyan eset áll fenn, hogy ezzel az algoritmussal átugorhatjuk a jobb lehetőségeket, akkor a visszafele haladóval ezeket megkeressük. A két algoritmust a következőképpen egyesítjük:

Az előre haladó vizsgálatkor még egy lépést kell beiktatni a szakaszok vizsgálatába. Azon kívül, hogy a szakaszon belül kezdődő, a végpontot következő szakasznál jobban túllépő szakaszokat figyelembe vettük, azt is meg kell vizsgálni, hogy az  $m_i$  szakaszon belül kezdődik-e olyan szakasz, amely hosszabb, mint  $h_i$  ( $m_i$  hossza), amely viszont a szakasz végpontját nem lépi túl jobban, mint az utána következő szakasz. Ha van, ugyanígy megvizsgáljuk ezt a szakaszt is, egészen addig, amíg találunk ennek a feltételnek megfelelő szakaszt. Ha már nem, akkor megjegyezzük azt a helyet, ahonnan a vizsgálatot kezdtük, valamint ahol az utolsó ilyen hosszabb belső szakasz véget ért, és ettől a végső ponttól kezdve végzünk vizsgálatot visszafelé haladva a megjegyzett kezdőpontig a fent leírt algoritmussal. A két algoritmus minden esetben megtalálja a legjobb fedést biztosító szakaszokat. Az előre haladó algoritmus úgy választ ki kettőt 3 szakasz közül, hogy az első kezdőpontja ( $s_i$ ) és a második végpontja ( $f_i$ ) között maximális legyen a távolság, a két szakasz közül pedig a hosszabbnak a hosszát tartja meg, a másikat pedig csökkenti. Ez a választási módszer elfedi azt az esetet, amikor az  $s_i$  és az  $f_i$  között van egy  $h_j$  hosszúságú szakasz, ahol  $s_j > s_i$  és  $f_j < f_i$  és  $h_j \geq h_i$ . Ezért ilyenkor a keresést a hátrafele haladó algoritmussal kell elvégezni  $s_i + h_j$  ponttól. Az  $i$ -edik pozíciótól hátrafele haladó algoritmus az olyan  $m_j$  szakaszok közül, ahol  $j < i$ ,  $s_j < i$  és  $f_j \geq i$ , azt választja ki, ahol  $s_j - i$  maximális: ez lesz  $m_k$  szakasz. A hátrafele keresés akkor nem találja meg az optimális megoldást, ha van olyan  $m_l$  szakasz, hogy  $l < k$ , és  $h_l > h_k$ , de  $s_k < f_l < i$ . De mivel a hátrafele haladó algoritmust

csak akkor alkalmazzuk, ha a később kezdődő szakasz hosszabb vagy egyenlő, mint az előbb kezdődő, tehát minden  $j < i$ -re,  $h_j \leq h_i$ , tehát nem fordulhat elő olyan  $m_i$ , hogy  $l < k$  és  $h_i > h_k$ . Az algoritmus tehát a büntetési képlet szempontjából mindig az optimális felbontást találja meg.

Az algoritmus leírása:

```
Function FelbontasHatra(intervallum, első index, utolsó index, szakaszok)
{
    az intervallum utolsó indexétől kezdjük a vizsgálatot
    ciklus amíg (nem értük el az első indexet) /* az elejéig megyünk a mondatnak */
    {
        ciklussal megkeressük az aktuális elemtől balra legtávolabbi szakaszt,
            amely még lefedi az aktuális elemet is,
            és megjegyezzük ennek pozícióját

        ha volt ilyen szakasz, amely lefedi az aktuális elemet
        {
            a szakasz nem átfedő részét felvesszük a szakaszok közé
            a szakaszok tömbben továbblépünk
            az intervallumban a következő elemre ugrunk: az előbb megjegyzett szakasz előtti elem
        }
        különben
            az intervallumban eggyel előző elemre lépünk
    }
    szakaszok tömb elemeinek sorrendjének megfordítása
    visszatérési érték: a szakaszok tömb elemeinek száma
}
```

A `FelbontasKeszit` függvény elkészíti a legjobb felbontást, amellyel a `MondatFelbontas` tömböt tölti fel. A függvény visszatérési értéke a felbontásban szereplő szakaszok száma.

```
Function FelbontasKeszit(intervallum, felbontások, intervallum hossza)
{
    az intervallum első elemétől indulunk
    ciklus amíg nem értük el az intervallum végét
    {
        ha az aktuális elemtől nem kezdődik szakasz
```

```

{
    az elemet átugorjuk, továbblépünk az intervallumban
}
különben ha nem túllépő szakaszrész vizsgálunk és az aktuális
szakasz hossza 1
{
    az aktuális szakaszt felvesszük a felbontásba következő
    elemnek
}
különben ha vége van az intervallumnak
{
    az aktuális szakaszt felvesszük a felbontások közé
    az intervallum végére lépünk
    a felbontások tömbben tovább lépünk
}
különben ha az aktuális szakasz túllép az intervallum végén
{ /* gyakorlatban nem fordulhat elő */
    az aktuális szakasz intervallumon belüli részét felvesszük a
    felbontások közé
    az intervallum végére lépünk
}
különben
{
    ha nem túllépő szakaszrész vizsgálunk
    {
        megjegyezzük az aktuális szakasz hosszát: felvesszük a
        felbontások közé
    }
    különben /* túllépő szakaszrész vizsgálunk */
    {
        az aktuális szakasz hossza a felbontások tömbbe
        utoljára felvett hossz lesz
        /* mert ezt a túllépő szakaszt már felvettük a
        felbontások közé */
    }
    a szakaszon belül az első elemmel kezdünk
    ciklussal megkeressük az aktuális szakaszon belül kezdődő
    szakaszok közül azt, amelyik legjobban és a megfelelő
    mértékben túllépi a szakasz végpontját,
    és megjegyezzük a túllépő (nem átfedő) rész hosszát és
    pozícióját
}
ha volt az aktuális szakaszban megfelelően túllépő szakasz

```

a két szakasz közül az marad meg a felbontások között, amelyik hosszabb volt, a másiknak az a része kerül a felbontások közé, amely nem fed át: ha az első szakasz volt hosszabb, a felbontások között marad az eredeti értéke, ha ez a rövidebb, akkor a felbontások közé már felvett értékét csökkentjük

```

különben ha nem volt túllépő szakasz
{ /* ha nem volt megfelelően túllépő szakasz, meg kell vizsgálni,
  hogy van-e az aktuális szakasznál hosszabb belső szakasz */
  ciklussal addig haladunk végig az intervallum szakaszain,
    amíg a szakaszban találunk vele egyenlő hosszúságú
    vagy hosszabb belső szakaszt,
    és megjegyezzük az utolsó ilyen szakasz végét
  ha volt ilyen hosszabb, belső szakasz
  {
    a megjegyzett végpozíciótól a megjegyzett
    kezdőpozícióig elvégezzük a visszafelé haladó
    vizsgálatot:
    FelbontasHatra(intervallum, kezdőpozíció, végpozíció,
    szakaszok)
    az így megkapott szakaszokat felvesszük a felbontások
    közé
  }
  különben /* ha nem volt túllépő szakasz vagy belső, hosszabb
  szakasz */
  {
    az aktuális szakasz hosszát felvesszük a felbontások
    közé
  }
}
visszatérési érték: a felbontásban szereplő szakaszok száma
}

```

Az algoritmus C jellegű pszeudokódját a 11.4 melléklet tartalmazza.

#### 5.6.2.5. A képlet módosítása

A büntetési képlet módosítására is szükség van. Az internetes értékelő módszerben ugyanis maximum 5 hosszúságú n-gramokra, valamint a teljes szegmensre keresünk rá. Emiatt a leghosszabb szakasz legfeljebb 5 szó hosszúságú lehet. Így a képletet úgy módosítjuk, hogy a nevezőben a maximális szakaszhossz



helyett több szakasz hosszát is felhasználjuk. Azt, hogy hány darab ilyen maximális hosszúságú szakaszt használunk fel, aszerint határozzuk meg, hogy ahány darab 5 hosszúságú szakasz férne el egy szegmensben (átfedés nélkül). A hosszt nevezzük  $mx$ -nek, amely az  $i$ -edik maximális hosszúságú szakaszt jelöli. Az  $i$  maximális értékét (egész szám,  $mx$ -szel jelöljük) egészszószttással határozzuk meg:

$$mx = h / 5$$

A büntetés tehát a következő lesz:

$$B = 0,5 \cdot \left( \frac{s + 2n}{2h + \sum_{i=1}^{mx} max_i + 2m} \right)^3$$

A METEOR érték tehát egy szegmensre, az új büntetési képlet használatával:

$$M = F_{mean} \cdot (1 - B)$$

A harmonikus átlag kiszámításához a paraméterek kezdeti, 2005-ös értékét használjuk, tehát  $\alpha = 0,9$ .  $P$  a pontosság a fent bemutatott képlettel, az  $R$  a fedés, szintén a fent leírt képlettel.

$$F_{mean} = \frac{P \cdot R}{0,9 \cdot P + 0,1 \cdot R}$$

#### 5.6.2.6. Unigramok kezelése

Szintén felmerül az unigramok kezelésének problémája is. Hiszen mivel nem referenciafordítást használunk, annak a jelentősége, hogy egyetlen szóra van találat a hatalmas internetes korpuszban, szinte elhanyagolható. Ha emberi szövegek minősítéséről van szó, akkor ennek egyáltalán nincs jelentősége, és még a gépi fordításban is nagyon kevés az olyan szó, amelyre nincs találat. Felmerül tehát az a lehetőség, hogy az unigramokat teljesen figyelmen kívül hagyjuk. Így viszont ugyanazt a büntetést fogja kapni az a szegmens is, ahol minden unigramra van találat, de szópárookra már nincs, és az is, ahol egyetlen unigramra sincs

találat. Emiatt tehát mégis érdemes figyelembe venni az unigramokat is. A másik indok arra, hogy az unigramokat is szakasznak vegyük, az, hogy egyrészt a gépi fordítás gyakran egészen érdekes szóösszetételeket hoz létre, ezért az unigramok megtalálásának – vagy inkább ha nincs rájuk találat – is nagy jelentősége van. Harmadrészt a szakaszok megállapításánál az átfedések miatt 1 hosszúnak kell venni olyan n-gramokat is, amelyek a valóságban ennél hosszabbak. A három érv alapján az unigramokat is ugyanolyan szakasznak vesszük, mint a hosszabb szakaszokat.

## 5.7. Eredmények

Az új, módosított METEOR módszert teszteltük a korábban is már többször használt szövegekkel, nyers gépi fordításukkal, a minimálisan utószerkesztett változatukkal és a teljesen utószerkesztett változatukkal. Az általunk kidolgozott módszert is teszteltük ugyanezekkel a szövegekkel, és az eredményeket összehasonlítottuk. Nézzük meg a szövegek átlagos eredményét „szövegtípusokra” lebontva.

Szöveg jellege	Saját értékelés	METEOR eredmény
Emberi fordítások:	0,872	0,892
Gépi fordítások:	0,358	0,850
Minimálisan utószerkesztett gépi fordítások:	0,525	0,880

18. táblázat

Az így módosított METEOR módszer eredményeivel nem vagyunk elégedettek, mivel nem tükrözik kellőképpen a különbséget a szövegek minőségében. Ennek oka az, hogy a képlethez szükséges fedés (recall) nem a valódi fedést tükrözi. Ez egy kényszerűségből létrehozott mérőszám, amelynek azonban nagyon nagy jelentősége van, hiszen 9-szeres súllyal számít. Ezért a mi esetünkben, az internetes korpusz használatával nem tartjuk megfelelőnek a METEOR módszer használatát a minőség méréséhez. A módosított büntetést és a legjobb felbontást megkereső algoritmust viszont a kutatás pozitív eredményének tartjuk.

## 5.8. Felmerülő kérdések

Az internetes mérési módszert illetően felmerülhetnek kérdések a korpuszal kapcsolatban. Nem félrevezető, ha egy szöveg többször is szerepel az interneten, és így a keresésre több eredményt kapunk? Nem tartjuk félrevezetőnek, mert abból a feltételezésből indulunk ki, ha egy szöveg többször szerepel, akkor ez azt jelenti, hogy több embernek tetszik, tehát több ember tartja elfogadhatónak. Másik kérdés lehet, hogy vajon hogyan befolyásolja az eredményt, ha egy szöveg csak „viccből”, éppen az elfogadhatatlanság szemléltetésére szerepel a korpuszban? Feltételezésünk szerint az ilyen szövegek a teljes korpusznak csak elhanyagolható részét alkotják, valamint annak a valószínűsége is elhanyagolhatóan csekély, hogy pontosan ugyanazok az n-gramok szerepelnek ezekben a szövegekben és az értékelendő szövegekben.

Az értékelési módszer egy másik érdekessége, hogy mivel a korpusz folyamatosan változik, az értékelés eredménye is változni fog. Az értékelés szempontjából a legfontosabb az összehasonlítás, a javulás mértékének megállapítása, emiatt fontos, hogy az összehasonlítandó szövegeket egymáshoz közeli időpontokban értékeljük. Valamint feltételezhetjük, hogy egyrészt a korpusz bővül, másrészt pedig elfogadható szövegekkel bővül, tehát a nem elfogadható sorozatokra továbbra sem lesz találat. Így az eredmény megbízhatóságát a korpusz változása nem befolyásolja.

## 5.9. Az eredmények összefoglalása

A fejezetben bemutatunk egy új értékelési módszert a géppel fordított szövegek elfogadhatóságának megállapítására. A létrehozott értékelési módszer egy  $[0;1]$  intervallumba eső számmal értékeli a szövegeket úgy, hogy az interneten található magyar szövegeket használja ehhez referenciaként. A módszerhez nem szükséges a költséges és időigényes emberi fordítói munka, és objektíven vizsgálja az elfogadhatóságot, amely egy nagyon szubjektív értékelési szempont. A különböző kísérleti értékelések és a kérdőíves felmérés után az értékelési módszerről

bebizonyítottuk, hogy használható és megbízható. Megfelel a Hovy et al. által megadott kritériumoknak is, amelyekkel egy automatikus értékelő módszernek rendelkeznie kell (Hovy et al. 2002c). A mérési indexnek van alsó és felső határa, 0 és 1. A 0 a lehető legrosszabb minőséget jelenti, amit bizonyítottunk is a mesterségesen elkészített szövegekkel. A monotonitás követelménye csak az emberi ítélettel összehasonlítva tesztelhető. Az emberi ítélet és az automatikus módszer Spearman korrelációs együtthatója 0,96 volt, ami nagyon jó korrelációt jelent, tehát a monotonitás követelménye teljesül.

A módszer hátránya, hogy az interneten levő tartalom folyton változik, tehát ugyanarra a szövegre sem kapunk kétszer egyforma eredményt. (Itt meg kell jegyezni, hogy az egyéb, referenciafordításon alapuló módszerek eredménye is különböző lesz ugyanarra a szövegre, ha más fordítók szövegeit használjuk referenciaként.) Ezért fontos, hogy ha a módszert szövegek összehasonlításra használjuk, akkor a mérések legrovidebb időn belül kell elvégezni. A módszer másik hátránya az, hogy nagyon erősen a keresőoldalak felépítésére, szerkezetére támaszkodik, így ezeket az oldalakat folyamatosan figyelni kell, és a változásoknak megfelelően kell a szoftvert frissíteni.

A kutatás eredményeképpen a METEOR módszer töredezettségi büntetési értéke helyett sikerült egy új képlettel jobb büntetési módszert találni, amely az értékelendő mondatok lefedettségét nagyobb súllyal veszi figyelembe. A módszer tesztelésének másik eredménye egy kis műveletigényű algoritmus, amellyel megtaláljuk egy intervallum optimális lefedését a megadott szakaszokkal úgy, hogy a lehető legnagyobb fedettség megvalósuljon a lehető leghosszabb szakaszokkal, tehát a kapott lefedési felbontás új büntetési értéke a legkisebb lesz. A mérési eredmények alapján ebben a formában a METEOR módszert nem tartjuk megfelelőnek a szövegek értékeléséhez az internetes korpusz alapján, mindazonáltal a büntetési módszert és az új algoritmust új és pozitív eredményként értékeljük.

## 6. A gépi fordítás javításának lehetőségei

### 6.1. A gépi fordítás javítása

A géppel fordított szövegek javítására több lehetőségünk is van, a fordítási folyamat különböző szakaszaiban. A javítási módszereket aszerint csoportosítjuk, hogy a folyamat melyik szakaszában történnek. A szakaszokon itt nem a kódolás, transzfer, dekódolás hármast értjük, hanem a CNY-i szöveg létrehozásától a FNY-i szöveg létrehozásáig zajló különböző tevékenységeket. Ezt a folyamatot három részre oszthatjuk, a középső szakasz maga a fordítás.

#### 6.1.1. A fordítás előtt

Már a tényleges fordítás előtt is befolyásolhatjuk a CNY-i szöveg minőségét. Mivel a FNY-i szöveg jellege, nehézsége hatással van a fordított szövegre, természetes lehetőségnek tűnik, hogy ha a FNY-i szöveget egyszerűsítjük, akkor a fordítás minősége jobb lesz. Az ilyen, egyszerűsített nyelven készült szöveget hívjuk kontrollált vagy korlátozott nyelvű szövegnek (controlled language text). A fordítandó szöveg eleve készülhet ezen a korlátozott nyelven, vagy utólag is lehet a FNY-i szöveget módosítani, ekkor előszerkesztésről beszélünk. (Somers 1998: 137)

Az, hogy milyen tényezők teszik nehezzé vagy egyszerűbbé a fordítást, fordítórendszerenként különbözik, tehát ezek a tényezők a fordítórendszer jellemzőitől, fordítási módszereitől függenek. Ez azt jelenti, hogy a rendszer sajátosságait figyelembe véve létezik egy sablon és szabályok, amelyek megadják, hogy a korlátozott nyelven milyen szerkezeteket, lexikai elemeket, formátumokat lehet használni. Ezek használatával a rendszer a szöveget nagy valószínűséggel jó minőségben le tudja fordítani. Az előszerkesztés lehetőségeit, a kontrollált nyelvet és ezek hatását a gépi fordítás minőségét többek közt Allen és Hogan, Mitamura, Aikawa et al. valamint Fiedere és O'Brien mutatja be. (Allen és Hogan 2000: 62-71, Mitamura 1999: 46-52, Aikawa et al. 2007: 1-7, Fiederer és O'Brien 2009: 52-74)

### **6.1.2. A fordítás alatt**

A fordítás alatti javításon két teljesen különböző folyamatot érthetünk. Bizonyos rendszerek lehetővé teszik, hogy a fordítás folyamata közben módosítsuk a szöveget, válasszunk az alternatívák közül, vagy javítsunk. Ezt a fordítást hívjuk interaktív fordításnak, de ez nem nevezhető teljesen automatikus fordításnak, ezért bővebben nem foglalkozunk vele. (Somers 1998: 137)

A másik folyamat magának a rendszernek a javítása. A legnagyobb hangsúly a kutatásokban ezen van, hiszen a cél a minél jobb minőségű fordítás létrehozása. Ebben a folyamatban viszont egyrészt nagyon kevesen tudnak részt venni, csak közvetlenül a rendszer fejlesztői, másrészt pedig ez a folyamat viszonylag lassú. A felhasználó viszont nem várhatja meg, amíg a rendszert tökéletesítik. A fordítás előtti és utáni javítási folyamatot egy széles réteg el tudja végezni, ráadásul belátható időn belül.

### **6.1.3. A fordítás után**

Ha már készen van a nyers fordítás, ennek javítását utószerkesztésnek hívjuk. Allen definíciója alapján az utószerkesztés az a folyamat, amikor az utószerkesztő személy szerkeszti, módosítja vagy javítja a géppel készített nyers fordítást (Allen 2003: 297). Bár Allen csak „emberi” utószerkesztést említ, számos kísérlet folyik az utószerkesztés automatikussá tételére. Ennek motivációja teljesen érthető, hiszen a gépi fordítás előnye pont a sebesség és az alacsony költség, amely nagyrészt elveszik az időigényes utószerkesztéssel.

Az utó- és előszerkesztésnek is vannak előnyei és hátrányai. Az előszerkesztés annyiban könnyebb művelet, hogy egy érthető, jól olvasható, ember által létrehozott, természetes nyelvű szöveget kell egyszerűsíteni. Utószerkesztéskor a gép által létrehozott mesterséges nyelvű szöveget kell javítani, tehát gyakran az eredeti szövegre kell támaszkodni, a nyers szövegen sokszor nehéz kiigazodni, és előfordulhat, hogy egyes mondatokat teljesen át kell fogalmazni. Előszerkesztés esetén viszont, annak ellenére, hogy a fordítás várhatóan jobb lesz, mégsem

egészen biztos, hogy nem fog hibákat tartalmazni. Utószerkesztéskor már a kimenetet javítjuk, tehát az eredmény megjósolható.

Az olyan fordítórendszereknél, amelyek nem területspecifikusak, valamint a célzott közönség és felhasználási terület széles körű, tehát a felhasználótól nem lehet elvárni a bemeneti szöveg módosítását, vagy a szöveg módosítására nincs is lehetőség (weboldalak esetében), az utószerkesztésre kell inkább koncentrálni az előszerkesztés vagy korlátozott nyelv használata helyett.

#### **6.1.4. Fordítás vagy utószerkesztés?**

Az utószerkesztés időigényes művelet, tulajdonképpen tehernek mondható, mert a fordítási folyamathoz képest kevésbé tekinthető alkotó tevékenységnek, nagy koncentrációt igényel, aprólékos, és rutinszerűnek, unalmasnak tűnik. Emiatt sokan foglalkoznak azzal a kérdéssel, hogy vajon a nyers gépi fordítás utószerkesztése vagy az emberi fordítás adja az elfogadhatóbb, hasznosabb eredményt, valamint hogy a fordítás vagy az utószerkesztés művelete kedveltebb, kevésbé fáradtságos. Azt is vizsgálják, hogy az utószerkesztés különböző fokozatai milyen eredményeket hoznak az említett szempontok szerint. Tehát a kérdés az, hogy érdemes-e egyáltalán az utószerkesztéssel foglalkozni?

A korábban említett vizsgálatokból kiderült, hogy az angol-magyar nyers gépi fordítás nem elfogadható, tehát a jelenlegi állapotban vagy csak nagyon korlátozottan lehet használni, vagy javítani kell rajta.

Picken és Melby alapján Hatim és Mason azt a véleményt képviseli, hogy a fordítók: inkább fordítanak, mivel ez kevésbé fárasztó, valamint a fordítók kevésbé elégedettek az utószerkesztési munkával, mint a fordítással (Hatim és Mason 1990: 24). Ezzel szemben a METEO rendszer tapasztalatai alapján Arnold et al. pont az ellenkezőjét állítja (Arnold 1994: 158). A fordítók sokkal elégedettebbek voltak a javítással, mint a szinte teljesen egyforma szövegek fordításával. Ezt támasztja alá Hutchins és Somers is (Hutchins és Somers 1992: 208), valamint

O'Brien is kimutatta, hogy az utószerkesztési erőfeszítés kisebb, mint a fordításra szánt erőfeszítés (O'Brien 2007: 52-74).

Church és Hovy azt állapította meg, hogy míg a fordítók kevésbé fogadják el a gépi fordítást, mint az egyéb felhasználók, érdekes módon a gyors utószerkesztést szívesen vállalják, és érdekes munkának tartják. A vélemények szubjektív voltát bizonyítja az, hogy a fordítók pozitívabban vélekednek ugyanarról a feladatról, ha tudják, hogy ha akarnak, visszatérhetnek a régi, hagyományos módszerhez (azaz a fordításhoz) (Church és Hovy 1993: 10). A hajlandóság az utószerkesztésre és a munkáról való vélemény nagyon szubjektív, és nagyon sok tényezőtől függ, így nem lehet általános, mindenkre vonatkoztatható megállapításokat tenni. Ezek a szubjektív vélemények viszont változtathatóak anélkül, hogy maga az utószerkesztési feladat változna.

Az utószerkesztés vizsgálatakor figyelembe kell venni azt is, hogy az utószerkesztésnek is vannak szintjei, nem feltétlenül kell a szöveget annyira kijavítani, hogy az emberi fordítással azonos minőségű legyen. Az utószerkesztést általában két, esetleg három fokozatra különítjük el: gyors és teljes utószerkesztésre, valamint Allen és Hogan említi a minimális utószerkesztést is. Löffler-Laurian két kategóriát ad meg, a gyors és a teljes utószerkesztést de nem adja meg pontosan a határvonalat a kettő között (Krings és Koby 2001: 45). A gyors és a minimális utószerkesztés között elmosódnak a határok, és nem definiált, hogy mi köztük a különbség (Allen 2001: 27, Allen és Hogan 2000: 65), így mi a két kategóriát egynek tekintjük, és ezentúl minimális utószerkesztésként hivatkozunk rá, mivel ez a terminus jobban kifejezi az utószerkesztés jellegét.

A gyakorlatban problémát jelent, hogy nincsenek általánosan elfogadott, tisztán definiált szabályok az utószerkesztéshez, főleg a minimális utószerkesztésre vonatkozóan (Allen és Hogan 2000: 64). Általánosan elfogadott, hogy a minimális utószerkesztéshez épphogy a szükséges javításokat kell elvégezni ahhoz, hogy a szöveg érthető és pontos legyen. Ez a szöveg még mindig közel van a gépi fordításhoz, inkább mesterséges nyelvű, mint természetes. A



teljes utószerkesztéshez pedig a szövegnek helyesnek kell lennie nyelvi szempontból, jól olvashatónak, érthetőnek, pontosnak kell lennie, tehát összességében a teljes utószerkesztéssel készült szöveg már ember által készített szöveg, és nem mesterséges nyelvű, még akkor is, ha elmarad a professzionális emberi fordítás minőségétől. Saját tapasztalatok alapján a minimális utószerkesztés könnyebb és gyorsabb feladat, mint a jó minőségű fordítás. A teljes utószerkesztés több energiát és időt igényel, és jelentősen többet, ha el akarunk szakadni a forrásnyelvi szövegtől, és nem akarjuk, hogy az eredmény „fordításízű”, idegenszerű legyen.

A problémát az jelenti, hogy ha nincsenek szabályok, útmutatók, akkor az általános tendencia az, hogy az elvárt minimális utószerkesztésnél általában jóval több javítást végeznek el a szerkesztők (Allen és Hogan 2000: 64). Saját tapasztalatunk is ezt támasztja alá. A következő fejezetben leírt vizsgálathoz szükség volt utószerkesztett szövegekre, az útmutatásban (ebben az esetben szándékosan) csak annyi szerepelt, hogy a szöveget épphogy annyira kell kijavítani, hogy érthető és olvasható legyen. A feladatot a különböző szerkesztők egészen másképp értelmezték, sokan inkább a teljes utószerkesztéshez közelebbi szövegek készítettek. Emiatt a gyakorlati utószerkesztési feladatoknál szükség van arra, hogy minimális javítandó kategóriákat állítsunk fel, amelyet a következő fejezetben fogunk látni.

Az utószerkesztési feladat megadásakor időkorlát megadásával is elkülöníthető a teljes és a minimális utószerkesztés, ahelyett, hogy a javítandó elemeket adnánk meg. Az időkorlattal rákényszeríthetjük a szerkesztőt, hogy a szükségesnél ne töltsön több időt egy szöveggel, de az egyénenként változó képességek és teljesítmény miatt lehetséges, hogy nem lesznek egészen egységesek a szövegek.

### 6.1.5. Az utószerkesztett szövegek elfogadhatósága

Bowker és Ehgoetz már korábban is említett felmérésében (Bowker és Ehgoetz 2005: 1-4) a résztvevőknek a három tényező: a minőség, költség és sebesség figyelembe vételével kellett választani az emberi fordítás, nyers gépi fordítás és utószerkesztett gépi fordítás között. A résztvevők 67,3%-a az utószerkesztett változatot választotta, ami azt jelenti, hogy az olvasók áldozatra is hajlandóak, ha a szöveg minőségén kívüli tényezőket is figyelembe kell venniük, hiszen az utószerkesztett szöveg minősége alulmarad az emberi fordításénak.

A gyors utószerkesztés védelmében hozhatjuk fel példának Roberts egy vizsgálatát arra vonatkozóan, hogy a különböző szinteken utószerkesztett szövegnek milyen az érthetősége. Az utószerkesztés két szintjét úgy határozza meg, hogy a teljes emberi fordítási időnek mekkora részét teszi ki. A gyors utószerkesztés az emberi fordítási idő harmadát veszi igénybe – ez nem feltétlenül minimális utószerkesztés –, a teljes pedig az a javítás, amelyet el lehet végezni a teljes emberi fordítási idő két harmada alatt. A kutatásból az derül ki, hogy a gyors szerkesztés 89%-os érthetőséget eredményez, a teljes emberi fordítás érthetőségéhez képest, a teljes utószerkesztés pedig 99%-os érthetőséget. Mivel az érthetőség nehezen mérhető, a vizsgálatban egy saját formulát alkalmaz a számszerűsítésére (Roberts 2007: 1-3).

Doyon et. al. utószerkesztéssel foglalkozó vizsgálatából (Doyon et. al. 2008: 2-7) szintén az derült ki, hogy a gyors utószerkesztés és a teljes utószerkesztés elfogadhatósága között nincs jelentős eltérés. Szintén vizsgálták a kézi és az automatikus utószerkesztés elfogadhatóságát. Míg a kézi utószerkesztés minden esetben elfogadható, a gépi (automatikus) utószerkesztés nem éri el egyetlen esetben sem az elfogadható szintet, bár az egyik utószerkesztő (WordPerfect) jelentősen javított az elfogadhatóságon.

A szakirodalomban olvasható vizsgálatok és bizonyítékok alapján, valamint a saját tapasztalatok és kísérletek után azt mondhatjuk, hogy érdemes az utószerkesztéssel foglalkozni.

## **6.2. Az automatikus utószerkesztés**

A gépi fordításnak a legnagyobb előnye a sebesség és az alacsony költség. Ha viszont embereket kell alkalmazni a javításhoz, ez jelentősen csökkenti ezeket az előnyöket. Ráadásul a fent említett vizsgálatok alapján az utószerkesztési munka megítélése is vegyes, nem mindenki által kedvelt. Ezek miatt nem csodálkozhatunk, hogy kísérletek folynak a folyamat automatizálására, amely visszaadná a gépi fordítás előnyeit.

Félig automatikus Guzmán módszere, amellyel az utószerkesztő személy reguláris kifejezések segítségével javítja ki az ismétlődő hibákat. (Guzmán 2007a: 49-52 és Guzmán 2007b (elektronikus)) Allen is egy olyan, kifejezetten az utószerkesztéshez készített programot ajánl, amely emberi javítást igényel, de a kiegészítő modulok és funkciók segítségével a minimális utószerkesztés (amelyet a szoftver minimális emberi beavatkozással, nagyrészt automatikusan végez) az emberi teljes fordítási időnek harmadát teszi ki. Sajnos arra vonatkozóan nem ad precíz, összehasonlítható adatokat, hogy a teljes utószerkesztés mennyire hatékony. Csak annyit tudunk meg, hogy annyi ideig tart, amíg egy ember újra legépelné a szöveget átlagosan 25-30 szó/perces gépelési sebességgel (Allen 2001: 1-6).

Knight és Chander már 1994-ben olyan utószerkesztő modulokért érvel, amelyek a különböző gépi fordítórendszerek között hordozhatóak. Azt írják, hogy "A general posteditor must be useful for improving texts produced by a wide variety of MT systems and non-native speakers and should operate equally across all domains." (Chander és Knight 1994: 779) „Egy általános utószerkesztőnek képesnek kell lennie többféle gépi fordítórendszerrel vagy nem anyanyelvi beszélő által létrehozott szöveget is kijavítani, valamint különböző területeken

egyformán kell működnie” (Chander és Knight 1994: 779). Valóban az lenne az ideális utószerkesztő modul, amely teljesen általánosan működik, de mi inkább azt az irányt választjuk, hogy egy rendszerre és egy nyelvre megvizsgáljuk a lehetőségeket, amelyeket később lehet általánosítani. Érdekes, hogy Chander és Knight megállapításával ellentmond a kísérleti utószerkesztő moduljuk, amely japán-angol nyelvű géppel fordított szövegekben pótolja a hiányzó névelőket. Ez a modul csak a japán-angol nyelvpárral használható, amit saját megállapításukkal is alátámasztanak: az „ez” és „az” mutató névmásokkal nem foglalkoznak, mivel ezek a japán nyelvben expliciten jelöltek. Ennyiben tehát nem általános ez az utószerkesztő modul sem. (Chander és Knight 1994: 779-783) Alapvetően kétféle utószerkesztő modult ajánlanak: az általánosat (general) és az adaptívat (adaptive). Az általánosról fent volt szó, az adaptív utószerkesztő pedig statisztikai gépi fordítási technikákat használhat a tanuláshoz. A következőkben szeretnénk bemutatni néhány, ezen az utóbbi módszeren alapuló kutatást.

A kézi utószerkesztést segítő modulok is már nagyban megkönnyítik az utószerkesztő munkáját, de a teljesen automatikus utószerkesztés irányában is sok kutatás folyik. Az automatikus utószerkesztés egyik fő irányvonala a statisztikai utószerkesztés. Ennek lényege minden kutatásban a statisztikai gépi fordításnál is használt módszer, a tanulás. A szoftvert minden esetben valamilyen tanító korpusszal tanítják, amely alapján utána módosítja a szöveget. Allen és Hogan az utószerkesztés automatizálása mellett arra koncentrált, hogy egy automatikus statisztikai utószerkesztő modullal kiszűrjék a különbségeket az utószerkesztett és az eredeti szövegek között, majd ezek segítségével utószerkesztési elveket definiáljanak. Az utószerkesztő modul szöveghármasokból tanul – eredeti szöveg, géppel fordított szöveg és annak utószerkesztett változata. A cél a leggyakoribb hibák kiszűrése – amelyeket ez a modul megtanul –, majd javítása (Allen és Hogan 2000: 63-70).

Dugast et. al. két hasonló elvű, de egymástól függetlenül, két utószerkesztő modullal végzett utószerkesztési kísérlet eredményeit foglalja össze. Mindkét

utószerkesztő modult, az angol–francia és francia–angol PORTAGE (Simard et al. 2007a: 508-515, Simard et al. 2007b: 203-206), és a német–angol és spanyol–angol Moses (Koehn et al. 2007: 177-180) a Systrannal, egy alapvetően szabályalapú rendszerrel kombinálták. A statisztikai utószerkesztő modulok a Systran nyers fordításából, a FNY-i szövegekből és a fordítások utószerkesztett változatait használja a tanulásra, majd újabb szövegek utószerkesztésére használható. A két kísérlet hasonló eredményt hozott: a BLEU index jelentősen javult (egyes nyelvek esetében akár 1 tizeddel), a szövegek folyékonysága javult, de az utószerkesztési műveletek a nagyobb mértékű javítás mellett rontottak is a szövegen (Dugast et al. 2007: 220-223). Számunkra fontos eredmény, hogy a kísérletben kapott utószerkesztett szövegek vizsgálatával javítási kategóriákat is sikerült felállítani, amelyek segítségével pontosan megállapították kategóriánként a javulások, az ekvivalens eredmények és a rontások arányát: 8 javulásra jutott 1 rontás, amely elfogadhatatlan.

Ez tanulmány, és az egy évvel későbbi, hasonló tanulmányuk fontos, mert utószerkesztési kategóriákat sorolnak fel (Dugast et al. 2008: 175-178). Ezek a kategóriák segítséget adhatnak a kutatásunkban megállapított javítások kategorizálására. Sajnos a kategóriák egy az egyben nem alkalmazhatók, a fordításban résztvevő nyelvek és a fordítórendszerek különbségei miatt, de hasznos őket figyelembe venni.

## 7. Utószerkesztési alapelvek és kategóriák

### 7.1. Az utószerkesztés szempontjai

A következőkben szeretnénk bemutatni az angolról magyarra géppel fordított nyers szövegek utószerkesztésére vonatkozó kutatásokat és ajánlásokat. Az utószerkesztett szövegek használatát, mint tanító korpuszt mi is ígéretesnek tartjuk. A korábban említett utószerkesztő modulok terület-specifikusak, azaz használatukkal a szabályalapú fordítórendszert adaptálhatjuk egy adott területhez. Az utószerkesztési vizsgálatunk alapján is az derült ki, hogy sok utószerkesztési műveletet érdemes terület-specifikusan megadni. Az általunk végzett kísérlet viszont általános, a szövegek szakterület szempontjából változatosak, szövegtípusuk és nyelvváltozatuk viszont állandó.

A szakirodalomban szereplő kutatások nagy többsége az angol és egyéb indoeurópai nyelvekre vonatkozik. A magyar nyelv gazdag toldalékolási rendszere ezekkel a nyelvekkel szemben véleményünk szerint más megközelítést követel a nyelvtechnológia részéről. Míg az izolálóbbr vagy flektáló nyelveknél a statisztikai megközelítés megfelelő lehet, a magyarban a formák gazdagsága miatt a statisztikai feldolgozáshoz jóval nagyobb korpuszra lenne szükség. Az automatikus utószerkesztés elvégzéséhez statisztikai alapon óriási javított korpuszt kellene felhasználni. Előzetes megfigyelések alapján ugyanis bár a hibák szisztematikusan jelennek meg, mégis annyira speciálisak, hogy célszerűbbnek látjuk a műveletek automatikus szűrése és kivonatolása elvégzése után a kézi szabályalkotást.

Az utószerkesztési elvek megalkotásához érdemes lehet figyelembe venni a pszicholingvisztika eredményeit. Pléh alapján „A mondatértés az a folyamat, ahogyan a szavakat szószervezetekké kapcsoljuk össze, a szervezetekhez mondattani funkciót rendelünk, illetőleg ezeket a szervezeteket szemantikailag is értelmezzük. A ténylegesen elhangzó információt számos, már ismert dologgal

kiegészít(het)jük” (Gósy 2005: 176). A mondatértés folyamatában háromféle elemzést végzünk, az első a propozicionális elemzés, a második az adott-új tagolás szerinti elemzés, a harmadik pedig a cselekvésérték elemzése. A propozicionális elemzés összetevőire bontja a mondatot, ezekhez nyelvtani funkciót rendel. Az elemzés során feltárjuk a szó szerkezeti és állítmányfüggvényi viszonyokat. A folyamat nyelvenként különböző, a gazdag morfológiájú nyelvekben az alaktani elemzésen van a hangsúly, a toldalékok szerepe nagy. Fontosságban ezt követi a szórend, majd az egyeztetés. Ezzel ellentétben a nem agglutináló nyelveknél a hallgató/olvasó számára a sorrendiség, a szórend és az egyeztetés az elsődleges támpont (Gósy 2005: 176). A megértéshez szükséges az adott és új információ felismerése és összekapcsolása. A megértés harmadik szakasza a cselekvésérték elemzése és azonosítása. A megértés harmadik szintjét már nem vizsgáljuk, de az adott és új információ azonosítása szempontjából érdekes kérdés, hogy milyen igény jelentkezik az olvasó részéről, hogy a mondatrészek sorrendje ennek megfelelő legyen.

Az utószerkesztési elvek meghatározásával tisztán elméleti alapon, az eddigi kutatási eredmények segítségével is talán érdemes lenne megpróbálkozni, de az eddigi kutatások alapján még nincsenek tisztán körvonalazódott kategóriák és egyértelműen elfogadott, a szükséges mértékben részletes megértési modellek. A pszicholingvisztika megértési modelljéből mindenképpen alapul vehetjük azt, hogy a legfontosabb a nyelvi hibák kijavítása. Ehhez elengedhetetlen cél az, hogy a szavakat szó szerkezetekké lehessen összekapcsolni, és fel lehessen deríteni az állítmányfüggvényi viszonyokat. Kérdés, hogy ezek után már sikerül-e a mondat adott és új információ tartalmát megállapítani. Vajon az olvasó a nem szokványos, idegenszerű szövegben annak ellenére is fel tudja ismerni az adott és új információt, hogy az elemek nem a magyar mondatoknál megszokott aktuális tagolás sorrendjében vannak? Hajlandó ehhez több erőfeszítést tenni? Ha nem, valószínűleg szükség lesz a mondat aktuális tagolásának, az elemek sorrendjének változtatására is.

A nemzetközi kutatások, elméletek és saját feltételezések alapján a leghasznosabbnak az empirikus, induktív elemzést ítéltük. Az elfogadhatóság szubjektivitása indokolja a vizsgálat módszerét, mégpedig azt, hogy több személyt kértünk meg a nyers gépi fordítások javítására, és a javítások alapján állapítjuk meg az eredményeket és a javítási kategóriákat. Feltételezzük ugyanis, hogy aki javítja, – saját maga számára – elfogadható minőségűre javítja a szöveget, még ha ez a minőség személyenként különbözik is.

## **7.2. A hibák értékelése**

### **7.2.1. Az utószerkesztés célja**

Az utószerkesztéshez szükséges a hibák feltérképezése, elemzése, majd javítása azért, hogy a szövegek érthetőek legyenek, az utószerkesztés szintjeinek megfelelően pedig az is, hogy csökkenjen a feldolgozási erőfeszítés. Pszicholingvisztikai kutatások alapján a hallgatók/olvasók jobban elfogadják a mondat jelentésbeli zavarait, mint a nyelvtani szerkezet hibáját (Gósy 2005: 131). Ezt bizonyítja a korábban bemutatott kérdőíves felmérés eredménye is. De az érthetőség is hatással van az elfogadhatóságra, és a szöveg elfogadhatósága is befolyásolja a feldolgozási erőfeszítést, ezen keresztül pedig az érthetőséget, tehát az elfogadhatóság és az érthetőség egymást befolyásolja. Szélsőséges esetekben egymás nélkül is jellemzőek lehetnek a szövegre: akkor is lehet érthető a szöveg, ha nem elfogadható, és lehet elfogadható akkor is, ha nem érthető. De ahhoz, hogy azt mondhassuk, a szöveg minősége jó, érthetőnek és elfogadhatónak is kell lennie.

Az elfogadhatóságnak tehát az egyik nagyon fontos aspektusa a nyelvi szerkezet. Ennek ellenére a megértési folyamatot a nyelvtanilag helyes, de mégis értelmetlen mondatok jobban akadályozhatják, mint a nyelvi szerkezet kisebb hibái. Gósy szerint viszont a megértés kulcsa az együttműködési elv funkcionálása a beszélő és a hallgató között, azaz a hallgató feltételezi, hogy a beszélő igazat



mond, amely megfelel a hallgató ismereteinek, a hallgató pedig szeretné ezt megérteni (Gósy 2005: 181, Reboul és Moeschler 2005: 53). De ha az olvasó nem tartja elfogadhatónak a szöveget a nyelvi hibák miatt, akkor ellenségesen fog hozzáállni, tehát nem lesz együttműködő. Így pedig nem lesz hajlandó erőfeszítéseket tenni a kontextuális hatás érdekében, akkor sem, ha ez kevesebb erőfeszítést jelentene, mint egy elfogadható, de nem érthető szöveg esetében. Ezért tehát a szöveg elfogadhatósága a gépi fordítás esetében rendkívül fontos. A fentiek alapján az elfogadhatóságot pedig a nyelvi hibák kijavításával növelhetjük. A gépi fordítás esetében pedig legtöbbször nem csak kis nyelvi hibákról van szó, és nem csak az elfogadhatósággal van probléma, hanem a tényleges megértéssel is. Bármennyire is együttműködő a hallgató a feldolgozás során, a megértés a nyelvi forma elemzése nélkül nem valósulhat meg. Emiatt tehát az érthetőséget akadályozó hibákat mindenképpen ki kell javítani.

Az érthetőséget leginkább akadályozó hibákat empirikus módszerrel állapítjuk meg. A hibaelemzéshez a gépi fordítás irodalmában az Arnold et al. által leírt módszert követjük annak ellenére, hogy nem feltétlenül javasolja. Erre három érvet sorol fel:

1. Az elemzés nagyon sok időt vesz igénybe.
2. Nincs következetes, jól használható súlyozási rendszer.
3. A géppel fordított szövegekben annyira sok a hiba, amelyek egymást befolyásolják, hogy az értékelésnek nem mindig van értelme (Arnold et al. 1994:166-167).

A három érvvel teljes mértékben egyetértünk, tapasztalataink alapján mindhárom igaz. Ennek ellenére úgy gondoljuk, az eredmény érdekében érdemes az elemzésre szánni az időt, az egymást befolyásoló hibákra pedig egyedi kategóriákat lehet létrehozni. A súlyozási rendszert pedig kérdőíves felmérés segítségével kívánjuk létrehozni.

### 7.2.2. Hibakategóriák és rendszerek

A hibaelemzés első nehézsége a hibakategóriák felállítása. Feltételezésünk szerint érdemes az elemzés előtt felállítani egy alapvető kategóriarendszert, amelyet majd az elemzés során bővíteni és alakítani kell, mivel az ember által létrehozott szövegekben előforduló hibák halmaza nem fedi a gépi, mesterséges nyelven létrehozott hibák halmazát, magyar gépi fordításokhoz pedig még nincs ilyen kategóriarendszer.

Annak ellenére, hogy egyetértünk Sager már többször említett megállapításával (Sager 1994:258), a géppel lefordított szöveget is emberek fogják olvasni, akiknek meg kell felelni. Ha ezt a szempontot vesszük figyelembe, akkor a géppel fordított szövegekre is megpróbálhatjuk alkalmazni az emberi fordítások értékelésére készült hibatipológiákat. A fordított szövegek értékelésének módja lehet analitikus és holisztikus. A felhasználó, azaz olvasó szempontjából a holisztikus értékelés a mérvadó, hiszen akkor fogja a géppel fordított szövegeket olvasni, ha ezek olvasása kevés erőfeszítést igényel. A szoftver fejlesztőinek szempontjából és az utószerkesztés miatt viszont szükség van az analitikus értékelésre, ezért az analitikus hibák tipológiáit tekintjük át röviden.

Hatim és Mason két fő csoportba osztja a hibákat, de csak az egyik csoportot nevezi tulajdonképpen hibának, azokat a helyzeteket, amelyekben a FNY-i és a CNY-i szövegben levő denotációs jelentés között jelentős indokolatlan eltérés van, vagy ha a CNY rendszere torzul, például nem megfelelő a helyesírás vagy a nyelvtan. Ezenkívül jelentkezhetnek hibák szövegszinten, a többi eset pedig mind fordítói választás eredménye, és nem nevezhető hibának (Hatim és Mason: 169). Mivel a gép a mondatot tekinti a legnagyobb egységnek, a szövegalkotásról egyáltalán nem lehet szó, a fennmaradó két kategória pedig nagyon kevés a gépi hibák rangsorolásához.

Sager öt hibakategóriája (jelentésfelcserélés, kihagyás, bővítés, eltérés, átalakítás) közül a gép nem hagy ki és nem bővít, legalábbis „emberi” értelemben,

így ez a felosztás céljainknak nem felel meg (Sager 1983: 101). Anthony Pym bináris (helyes - nem helyes) és nem-bináris (több jó megoldás) megkülönböztetése sem elég precíz jelen céljainkra (Pym 1992: 282-283). A funkcionalista fordítástudomány megközelítése szerint – egy kifejezést csak a fordítás céljának és a fordítói stratégiának figyelembevételével tekinthetünk adekvátnak vagy inadekvátnak – az a kérdés merül fel, hogy vajon a teljesen automatikus gépi fordítás lehet-e bármilyen szituációban adekvát. Ha nem lehet, akkor milyen javításokat kell rajta elvégezni, hogy az olvasó annak fogadja el? (Dróth 2001: 58) A jelen kísérletben még nem fogunk tudni választ adni a fenti kérdésre. A cél egyelőre az, hogy kiderítsük, a fordítás mennyire érthető a CNY-i olvasó számára (aki tudja, hogy a szöveget gép fordította). Newmark kategorizálása (Newmark 1993: 29-30) szerint súlyosság alapján kétféle, analitikus értelemben vett fordítási hiba van: a félrevezető és árnyalatnyi hiba. Félrevezető hibák lehetnek referenciálisak vagy nyelviak, árnyalatnyinak pedig a stilisztikai és lexikai hibákat tekintjük. Nagyon sok, nagyon súlyos hiba esetén ez a felosztás megfelelő lehet, mivel a hibák többé-kevésbé egyértelműen besorolhatóak a kategóriákba. Mivel ezeket a kategóriákat jó kiindulási pontnak tartjuk, ebből a négy hibatípusból indulunk ki, majd ezeket a kategóriákat finomítjuk tovább. A célunk annak megállapítása, hogy melyik hibatípus gátolja leginkább a szövegértést.

### **7.2.3. Eredmények és elemzés**

A kezdeti hibakategóriák felállítása után következik maga a hibaelemzés. Ezzel kialakul egy súlyozatlan kategórialista. A folyamat következő lépése a súlyok megállapítása. Ehhez Arnold et al. második szempontját vettük figyelembe, azaz a hibákat a szöveg érthetőségére gyakorolt hatásuk alapján súlyoztuk (Arnold et al. 1994: 166). Annak elkerülésére, hogy az érthetőséget egyetlen, szubjektív olvasó határozza meg, több olvasót kérdeztünk meg, és a súlyokat a kapott válaszok alapján állapítottuk meg. Az elemzés végén tehát kialakult egy súlyozási módszer

és egy kategóriarendszer, amelybe a hibákat besorolhatjuk. A négy kezdeti hibakategória feltételezett súlyossági sorrendje:

1. referenciális hiba
2. nyelvi hiba
3. lexikai hiba
4. stilisztikai hiba

Az is feltételezhető, hogy ezek a kategóriák nem lesznek elegendőek a gépi fordítás hibáinak besorolásához, de a kezdeti besoroláshoz jó alapot nyújtanak. Így tehát elemzés másik, a további kutatások szempontjából fontosabb célja az volt, hogy a szövegekben talált hibákat besoroljuk a kategóriákba, új kategóriákat hozunk létre, és a hibák között felállítsunk egy sorrendet az általuk okozott megértési nehézségek alapján.

A hibák kategorizálásához a korábban bemutatott kérdőíves vizsgálat anyagát használtuk. A kérdőív második része az volt, hogy a harmadik szövegben kiemelt részeket saját szavakkal meg kellett magyarázni. A kiemelt részek változó hosszúságúak, egy szótól a teljes (összetett) mondatig terjed. Az elemzéshez használt szöveg a 11.2 melléklet második oldalán található. Az ezekre a kérdésekre adott válaszok alapján állapítottuk meg a hibák súlyosságát. A kérdőívek elemzése előtt szükség volt a fordított szövegekben előforduló hibák előzetes kategorizálására, tehát harmadik szöveg hibái bekerültek a négy kategóriába.

- Referenciális hibának az tekinthető, ahol a szó, szó szerkezet vagy mondat a valóság egy másik részére vonatkozik a CNY-i szövegben, mint a FNY-i szövegben. (pl.: *szállítási gesztenye – közlekedési tréfa, történet*)
- Nyelvi hibák: ha a nyelvi rendszer torzul valahogyan, de ez a torzulás nem vezet referenciális hibához. (pl.: *minden útban – úton, hívás valami elhaladó helybelin*)
- Stilisztikai hiba: kicsit idegenszerű, furcsa mondatok, kifejezések. (pl. *van egy probléma, körökben száguldanak körbe*)

- Lexikai hiba: nem nyelvi és nem referenciális szó szintű hibák, rossz szóválasztás. (pl.: *tudod azt – ismered, öreg – régi*)

A besorolás után a kérdésekre adott válaszok alapján minden részlet kapott egy javítási indexet, amelyből kialakult a súlyossági sorrend. Az indexet hasonló elven határoztuk meg, mint az egyes részek érthetőségi indexét, de nem teljesen azonos módon.

A kérdések között szereplő hosszabb elemeket (pl. összetett mondatot) részekre kellett bontani, és úgy elemezni a válaszokat. Itt az elemzéskor nem az öt kategória (teljes válasz, hiányos válasz, részben jó, részben helytelen válasz, nem tudja, nagyrészt helytelen válasz, helytelen válasz) alapján soroltuk be a válaszokat, hanem egy [-5;+5] skálán osztályoztuk őket. Ennek indoka az, hogy az öt kategória nem volt elegendő, közöttük is kellett fokozatokat felállítani. Az 5-ös értelemszerűen a teljesen jó válasz, a -5 a teljesen rossz. Ha nem tudta a választ, ez itt is a skála közepén a 0 értéket kapta. A javítási index az ezekből az osztályzatokból kialakuló átlag.

Összesen 17 elem volt, amelyet osztályozni kellett, ezekre összesen 763 válasz érkezett. A következő táblázat mutatja, hogy mekkora a válaszok megoszlása az egyes osztályzatokban.

Osztályzat	Válaszok száma	Válaszok aránya
5	211	27,65%
4	74	9,70%
3	46	6,03%
2	33	4,33%
1	14	1,83%
0	50	6,55%
-1	4	0,52%
-2	10	1,31%
-3	229	30,01%
-4	4	0,52%
-5	88	11,53%

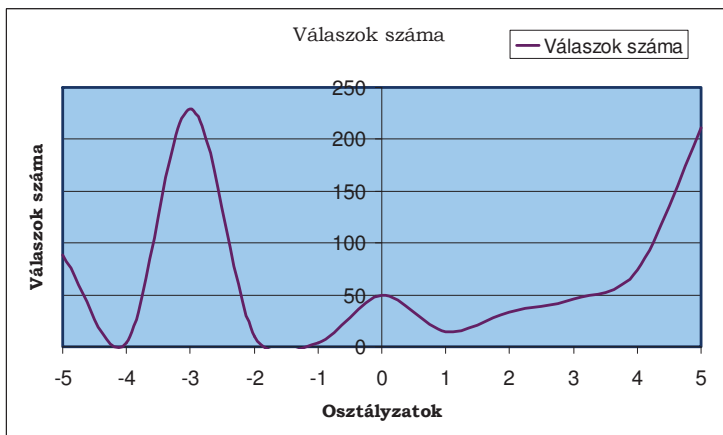
19. táblázat: A válaszok megoszlása osztályzat szerint

Itt láthatjuk ugyanezt a táblázatot a válaszok aránya szerinti csökkenő sorrendben.

Osztályzat	Válaszok száma	Válaszok aránya
-3	229	30,01%
5	211	27,65%
-5	88	11,53%
4	74	9,70%
0	50	6,55%
3	46	6,03%
2	33	4,33%
1	14	1,83%
-2	10	1,31%
-1	4	0,52%
-4	4	0,52%

20. táblázat: Értékek nagyság szerinti sorrendben

A táblázatból láthatjuk, hogy a listát a -3-as érték vezeti, utána pedig szinte alig lemaradva vannak az 5-ös válaszok. A -3 érték az a válasz, amelynek kicsit több, mint fele helytelen, azaz félig félreértett. Az 5-ös a teljesen jó válasz. A következő a sorban a teljesen rossz válasz, a -5-ös érték. Tehát az osztályzatok eloszlása a teljes skálán érdekes, mert két csúcspontja van szinte egyforma értékkel:



4. ábra Az osztályzatok megoszlása

A két csúcspont egyrészt azt jelentheti, hogy alapvetően kétféle típusú válaszadó volt, vagy azt, hogy a szövegrészek két szélsőséges csoportba sorolhatóak: vagy jól érthetőek, vagy félig félreérthetőek. Az eredmények vizsgálata után kiderül, hogy a második feltételezés igaz: a szövegrészeknél vagy a 4-es/5-ös vagy a -3-as válaszok aránya dominál. A legtöbb válasz a skála két végén helyezkedik el, a pozitív válaszok összesen 49,5%-át, a negatív válaszok pedig 43,9%-át teszik ki az összes válasznak. Ha a 0 értékű válaszokat a negatív értékekhez soroljuk, akkor a negatív értékek összege 50,4%. Azt mondhatjuk tehát, hogy a pozitív és negatív értékek aránya szinte egyenlő, alig 1% különbség van közöttük.

#### 7.2.4. Részletes hibakategóriák

Általánosan elmondható, hogy a szövegekben a nyelvi és referenciális hibák dominálnak, a stilisztikai hibák néhol el is vesznek ezek között. Várakozásaink szerint a hibák súlyosságának fent bemutatott sorrendje, és a szöveg többszöri elolvasása alapján a lexikai hiba kevésbé lesz súlyos, mint a nyelvi vagy a referenciális, így például a *sodródva megtalálják magukat* kifejezéstől azt vártuk, hogy jól érthető lesz, az *amiket tartalmaztak a nagy gondolkodásléggömbben, lebegni a feje fölött* tagmondat pedig nagyon rossz eredményt fog elérni. Ugyanígy a *Az hármát vágott az, hogy üveget röpített.* mondatnak a lexikai és nyelvi hiba miatt érthetőbbnek kellett volna lenni, mint a *szállítási gesztenye* kifejezésnek, és a sort folytathatnánk. Ehhez képest a kérdőívben adott válaszok meglepő eredményt hoztak.

A kérdőív eredményei alapján a négy hibakategória tehát nem bizonyult elegendőnek a sorrend megállapítására, így szükség volt ezek további finomítására, valamint néhány jelenséget a válaszok alapján át kellett sorolni másik kategóriába (ld. a táblázat utolsó oszlopát). Ilyen például a *sodródva megtalálják magukat*: lexikai és enyhe stilisztikai hibából (*eltévedve találják magukat*, azaz *eltévednek* vagy *rájönnek*, *hogy eltévedtek* lenne a „helyes”) referenciális lesz,

mivel a szöveg a kontextusa éppen az eltévedés, a *megtalálják magukat* szókapcsolat pedig a helyes út megtalálásának képzetét kelti az olvasóban.

Az elemzés után kialakított sorrend példákkal illusztrálva (legsúlyosabbtól a kevésbé súlyos hibákig):

1. referenciális hiba, kontextusba beleillik, nyelvileg helyes
  - több ezer (több tízezer helyett) (GF\_002)
2. mondat szintű referenciális és nyelvi hiba
  - melyik az, hogy róluk hol hiszik azt (GF\_003)
3. nyelvi és lexikai hiba: mondatrészek szerepe nem tiszta + szórend helytelen
  - Az hármát vágott az, hogy üveget röpített. (GF\_002)
4. lexikai szintű referenciális hiba + szó/kifejezés (hang)alakja teljesen eltér az odaillő kifejezésétől és jelentésmezőjük egyáltalán nem fed át
  - szállítási gesztenye (GF\_003)
5. kontextus miatt nyelvi hibából referenciális hiba
  - sodródva megtalálják magukat (rájönnek, hogy eltévedtek) (GF\_003)
6. szófajtévesztés: zárt lexikai halmazt érintő (névmások, kötőszavak, kérdőszavak stb.), kifejezetten a vonatkozó névmás vagy kötőszó helyett kérdő névmás használata
  - melyik az (amely), azelőtt (mielőtt) (GF\_003)
7. nyelvi hiba: helytelen toldalékolás (rag) miatt referenciális hiba
  - Acacia Avenue-nál Acacia Grove-ban találják magukat (helyett) (GF\_003)
8. nyelvi hiba: igeidő, mód tévesztése, elhagyása vagy ige helyett igenév használata
  - nem adni kifejezést annak (GF\_003)
9. referenciális hiba, de a szó alakja felidézheti az odaillő szót + a kontextusból felidézhető a kép vagy jelenet
  - ajkak annyira szorongtak egy vékony sorozatba (GF\_003)
10. lexikai hiba: nem létező szó vagy szóösszetétel használata
  - hajtó ülés (GF\_003)
11. stilisztikai hiba, tükörfordítás
  - bármilyen ilyen dolgok (GF\_003)
12. lexikai hiba, rossz szóválasztás
  - elveszék (eltéved helyett) (GF\_003)



13. nyelvi hiba: más szófaj, de a referenciális jelentés nem változik

- hívás (GF\_003)

14. nyelvi hiba: helytelen rag (igevonzat), ahol az igéből eleve kikövetkeztethető

- hívás valami elhaladó helybelin (GF\_003)

A következő táblázat a kérdőívben szereplő szövegrészeket és a javítási indexeket tartalmazza. Az utolsó oszlopban található a hiba sorszáma, amelyet a későbbiekben is használni fogunk.

Szövegrész	Index	Hiba típusa	Hiba leírása	Hiba sorszáma
mondott rendőrség által tíznek éves lenni több ezer	-2,92	ref. + nyelvi	mondatrészek szerepe nem tiszta + súlyos referenciális hiba, de nyelvtanilag helyes, kontextusba illik	1 + 2
És nem világosan elhivatottak... (melyik az, hogy róluk hol hiszik azt)	-2,6	ref. + nyelvi	értelmetlennek tűnő mondat	2
melyik az, hogy a fiam...	-2,5	ref.	szórend, mondatrészek szerepe nem tiszta	3
szállítási gesztenye	-1,9	ref.	referenciális jelentés teljesen más	4
megtalálják magukat	-1,55	nyelvi	RJ lehet ellentétes → R hiba	5
sodródva megtalálják magukat	-1,22	ref. + nyelvi	megtalálják magukat: ellentétes jelentés	5
Azáltal, hogy a kedves melyik időt üli meg az utasülésben, az ajkak annyira szorongtak egy vékony sorozatba nem adni kifejezést annak, hogy a szörnyű vélemények, amiket tartalmaztak a nagy	-1	ref. + nyelvi	addigra helyett Azáltal, hogy melyik időt RJ közel, de más jelentést ad a mondatnak	3+ 6 + 8 + 9

Szövegrész	Index	Hiba típusa	Hiba leírása	Hiba sorszáma
gondolkodásléggömbben, lebegni a feje fölött.				
sodródva	-0,8	ref.	RJ közel	5 + 12
Acacia Avenue-nál Acacia Grove-ban találják magukat:	-0,7	ref.	más jelentés a helytelen ragok miatt	7
nem adni kifejezést annak, hogy a szörnyű vélemények	0,1	nyelvi	szófaj helytelen (igenév), de RJ. ugyanaz + mód helytelen	8
ajkak annyira szorongtak egy vékony sorozatba	1	ref.	alak hasonlít + kép	9
tartalmaztak a nagy gondolkodásléggömbben, lebegni a feje fölött	1,2	nyelvi	nem létező, de RJ közel + kép	3 + 9
bíbor egy csúnya árnyéka	1,48	ref.	alak hasonlít	9
népkeszég	1,78	nyelvi	nem létező szó, de RJ közel	10/12
elismerjék, hogy ők bármilyen ilyen dolgok	1,86	stilisztikai	visszaautal a mondat korábbi részére (sodródva megtalálják magukat)	11
ez odajön ahhoz, hogy elvesszék	2,22	ref. + lexikai	RJ közel	9 + 12
hajtó ülés	2,98	nyelvi	nem létező szó, de RJ közel	10
elvesszék	3,3	lexikai	rossz szóválasztás	12
hívás valami elhaladó helybelin eligazítani őket:	4,4	nyelvi	más szófaj, RJ ugyanaz; helytelen rag	8 + 13 + 14

**21. táblázat: A szöveg hibái javítási index szerinti sorrendben**

Láthatjuk azt, hogy a kategóriák nagyon speciálisak, sokszor a külsőleg (szintaktikailag) ugyanabba a kategóriába eső hiba teljesen máshogy hat a megértésre, ha szemantikailag más kategóriába esik. Ez az elemzés ugyan csak egyetlen szöveg alapján készült, de feltételezhetően ez igaz a többi szövegre is,

még ha a pontos kategóriák nem is egészen azonosak. Ez a tény nagyon megnehezíti az automatikus hibajavítást. Valószínűleg ezen oknál fogva az utószerkesztés teljes automatizálásával próbálkozó kutatásokban általában a hibakategóriák szintaktikai, szerkezeti jellegűek, ezeken belül pedig szemantikailag nincsenek elkülönítve. (Ld. Dugast et. al. 2007: 222-223.) A súlyozás után látszik, hogy valóban a referenciális hiba a legsúlyosabb, de a nyelvi hibák is zavarják a megértést, sokszor keveredve a referenciális hibákkal. A négy kategória nem elkülöníthető, és a súlyosságuk sorrendje sem olyan egyértelmű, mint ahogy kezdetben feltételeztük.

Ki szeretnénk emelni azt, hogy a szövegben előforduló, nem lefordított elemeket nem vizsgáltuk, mert azok megértése az olvasó angol nyelvtudásától függ, tehát félrevezető eredményt ad, hiszen a gépi fordítás célközönsége éppen az adott forrásnyelven nem tudó olvasó. Ezeket a hibákat a megértésre gyakorolt hatásuktól függetlenül mindenképp ki kell javítani, tehát ezek akár vizsgálat nélkül is megkaphatják a -5-ös javítási indexet.

Megértés szempontjából a leghátrányosabb eset az, ha a kifejezés a szituációba beleillik, valamint nyelvilag is helyes, de a referenciális jelentés a FNY-ben és CNY-ben más, így az olvasó még csak észre sem veszi a hibát.

Nagyon súlyos az a helyzet is, ha a referenciális hiba mondat szinten jelentkezik, és nyelvi hiba is társul hozzá. Sajnos egy egyszerű helytelen lexikai választás a fordítórendszerben (például vonatkozó névmás vagy kötőszó helyett kérdő névmás) máris mondat szintű hibát eredményez, így súlyosan hátráltatja a megértést, ha pedig ehhez még nyelvi hiba is társul, a jelentés csak nagy nehézségek árán vagy egyáltalán nem kikövetkeztethető. Az ilyen esetek tulajdonképpen nyelvi hibák, de mivel teljesen megváltoztatják a mondat jelentését, ezeket a hibákat a referenciális hibák közé soroljuk.

Ezután következnek az olyan lexikai szintű referenciális hibák, amelyeknél a helytelen kifejezés alakja teljes mértékben eltér az oda illő kifejezésétől, valamint

jelentésmezőjükben sincs átfedés. (pl. szállítási gesztenye – közlekedési történet helyett).

Szintén viszonylag nagy problémát jelentenek az olyan nyelvi hibák (toldalékolás), ahol a helytelen toldalék referenciális hibát eredményez (*Acacia Avenue-nál Acacia Grove-ban találják magukat – Acacia Avenue helyett Acacia Grove-ban találják magukat*). Ezzel szemben az olyan toldalékolási hibák, amelyek nem eredményeznek referenciális hibát, például *hívás valami elhaladó helybelin*, szinte egyáltalán nem jelentenek problémát.

Érdekes jelenség az olyan referenciális hiba, ahol a szó alakja felidézi az odaillő szót, és a kontextusból felidézhető a kép vagy jelenet (*ajkak annyira szorongtak egy vékony sorozatba*). Az ilyen hibák nagyon kevés nehézséget okoznak, az olvasó ki tudja következtetni a jelentést. Megfelelő fordítás lenne az *annyira összeszorította az ajkát*, ahol a *szorong* szó jelentése ugyan más, mint a *szorít*, mégis egy a szótövük, a *vékony sorozat* pedig nem értelmes, mégis szinte magunk előtt látjuk az erősen összeszorított ajkakot.

Szinte egyáltalán nem okoz értési problémákat a lexikai és stilisztikai hiba, de ezek miatt a szöveg – esetleges komoly témája ellenére – derűltség forrása lehet. A szófajtévesztés (ige-főnév) (ha a referenciális jelentés megegyezik), és a helytelen ragok szintén csak kissé zavarják a megértést.

A pszicholingvisztikai kutatásokban még nincs egyértelmű válasz arra, hogy az előzetes tudás befolyásolja-e a nyelvtani típusú feldolgozást. A két szélsőséges irányvonal közül az egyik a formára alapoz, tehát a megértésben a kiindulás a szerkezetek azonosítása, majd utána következik a gyors tartalmi elemzés. A másik irány szerint a tartalom az alap, és a hallgató feltételezi, hogy a beszélő igazat mond, és meg akarja őt érteni. Gósy megállapítása szerint a kettő között csak kiindulási és hangsúlybeli különbségek vannak, a megértésben mindkettő jelen van (Gósy 2005: 181). A fenti elemzés is Pléh megfigyelését támasztja alá, mely szerint az asszociatív szerveződés a nyelv bármely része között létrejöhet (Pléh 1998 alapján Gósy 2005: 181). „A kezdet az, amit megértünk, azt kijelentésekbe

szervezzük, amelyekkel aztán számos (pl. asszociatív) műveletet végezhetünk az aktuális közlés pontos és egyértelmű dekódolása érdekében. A beszédmegértés során nyilvánvalóan minden lehetséges információt felhasználunk, amennyire lehet egyidejűleg” (Gósy 2005: 181). Ez adhat magyarázatot arra, hogy az *„ajkak ammyira szorongtak egy vékony sorozatba”* részlet érthetősége viszonylag jobb volt, mint amit az egyes elemek részletes elemzése után várnánk.

A javítási index alapján egyelőre így definiálhatjuk az abszolút minimális utószerkesztést: a minimális utószerkesztéshez a 0 és negatív javítási indexet kapott hibakategóriákat kell feltétlenül kijavítani. Ezek a hibakategóriák az 1-8-as számú hibakategóriák. A 8. fejezetben bemutatjuk az abszolút minimális utószerkesztés hatását az érthetőségre. A minimális utószerkesztéshez viszont majd további szigorításokat kell tennünk a nagyobb méretű utószerkesztett korpusz tanulságai alapján, mert a cél a teljes érthetőség.

### 7.2.5. Konklúzió

Az elemzési módszer előnye és hátránya is ugyanazokból a tényezőkből fakad. A kérdőíves módszer és a kézi elemzés lassú, viszont több alany részvételével hitelesebben megállapítható az érthetőség, a kézi elemzés pedig pontosabb, és a finomságokat is ki tudja szűrni. Sajnos azt is le kell szögezni, hogy a vizsgálatban részt vevő minta nagyon kicsi, és nem alkalmas minden hibakategória feltérképezésére. Ahhoz, hogy minden lehetséges előforduló hibát feltárjunk, elemezzünk, kategorizáljunk és súlyozzunk, sok szöveg kiértékelésére van szükség. A jelenlegi kérdőíves vizsgálati módszer ehhez nem célszerű, mivel ugyanazt a kérdőívet kell sok alanynak kitölteni. A felkért személyek türelme, ideje, energiája pedig véges, tehát hosszabb kérdőív már esetleg nem lenne hiteles. Ehelyett a következő részben leírt módszerrel sikerült több szöveget megvizsgálni, és ezek alapján további elemzéseket végezni.

A fent leírt kérdőíves vizsgálat jó alapot szolgáltat a hibakategóriák kezdeti listájának elkészítéséhez, valamint a hibák súlyosságának megállapításához.

### **7.2.6. Az eredmények összefoglalása**

Az előzőekben bemutatunk egy kérdőíves vizsgálatot, amelyben a szövegek érthetőségére koncentráltunk. Ezután az érthetőség alapján létrehoztunk részletes hibakategóriákat. A vizsgálatban részt vevő szövegek száma nagyon kicsi, emiatt messzemenő és általános következtetéseket nem szeretnénk levonni. A vizsgálat szándékaink szerint kiindulópontot jelent a további, nagyobb korpuszon végzett vizsgálathoz. Az érthetőség alapján megállapítottunk egy javítási indexet, amellyel a hibákat súlyosság szerint sorba rendeztük. A súlyossági sorrend támpontot ad az abszolút minimális utószerkesztéshez, valamint ezt összehasonlíthatjuk a későbbi felmérés eredményeként megkapott, az olvasók által végzett utószerkesztési műveletekkel.

A hibakategóriák és a súlyossági sorrend azt is megmutatja, hogy a szövegek hibáinak ellenére a látszólag teljesen elfogadhatatlan és érthetetlen szöveg is a vártnál nagyobb mértékben érthető.

## **7.3. Javítási minták**

### **7.3.1. Az utószerkesztési műveletek automatikus kivonatolása**

#### **7.3.1.1. A géppel fordított szövegek elemzése**

Ahhoz, hogy az utószerkesztési műveleteket, a hibákat, az utószerkesztési alapelveket meghatározzuk, a legcélszerűbb módszer a szövegek elemzése. A szövegek elemzésével megállapíthatóak a hibák, valamint az, hogy ezek mennyire rendszeresen és milyen környezetben fordulnak elő. Az elemzés történhet csak a géppel fordított szöveg vizsgálatával, de ez rendkívül sok munkát jelent az elemzést végző személynek. Ahhoz, hogy a megtaláljuk a rendszerességet, sok szöveg kell, amelyek elemzése sokáig tart. Az elemzőre háruló munkát csökkenthetjük úgy, hogy a szövegek utószerkesztését mások elvégzik, ezek után pedig csak a szövegeket kell összehasonlítani, és így elemezni a javításokat. Az

elemzést viszont így is kézzel kell végezni, és a szövegek összehasonlítása megfelelő eszközök nélkül szintén nagyon körülményes. Az elemzési idő rövidítésére célszerűnek látszik egy automatikus eszköz létrehozása, amely kiszűri a javítási műveleteket, és megfelelő formátumot biztosít az elemzés megkönnyítésére. A javítási műveletek elemzésére ezt a módszert követtük.

Az elemzéshez használt szövegek kiválasztásakor ugyanazokat az alapelveket tartottuk szem előtt, amelyeket a kérdőíves vizsgálathoz használt szövegek kiválasztásakor. Egyrészt olyan szövegekre volt szükség, amelyek érdekesek, és egy olvasásra nem túlságosan hosszúak. Úgy gondoltuk, hogy ennek megfelelően hosszuk maximum 500 szó legyen, hogy a szöveg hossza vagy unalmas témája a felkért javítókat ne riassza el a feladattól. A szövegek műfaja itt is internetes újságcikk, mivel a hírek mindenki számára érdekesek lehetnek, valamint újdonságot mondanak az olvasónak. A témák különbözőek, de igyekeztünk általános, szakzsargonteremtő cikket választani. A cikkek nyelve angol, minden cikk brit lapokból került ki. Összesen 58 db eredeti cikk van a mintában, amelyek közül néhánynak több javítása is szerepel a javított korpuszban, így a javított szövegek száma 61, az eredeti szövegekben a szegmensek száma 1316, a tokenek száma 26253, a javított szövegekben pedig tokenek száma 26134. A korpusz mérete a nyelvtechnológiai alkalmazásokéhoz képest nagyon kicsi, de pontosan ez teszi még lehetővé a kézi elemzést. Mivel itt nem statisztikai módszerről van szó, és a korpuszt egyelőre nem tanításra használjuk, a mérete elegendő ahhoz, hogy kiszűrjünk bizonyos tendenciákat, és további, nagyobb léptékű kutatásokhoz kiindulási pontot adjunk.

A szövegeket ezután a MetaMorpho ingyenes webfordítójával lefordítottuk magyarra. Csak ezt az egyetlen szabályalapú – pontosabban mintalapú – rendszert használtuk, mert egyrészt feltétlenül olyan rendszert akartunk vizsgálni, amely nyelvi tudásra támaszkodik, és nem statisztikai elemzést végez. Ilyen rendszer pedig angol–magyar fordítási irányban pedig csak ez az egy elérhető. A szintén weben elérhető Systran fordító a magyar nyelvre annyira rossz minőségű,

hogy ki kellett zárunk. A statisztikai rendszereket pedig azért zártuk ki, mert ott a rendszeresség esetleges, a háttérben levő korpusztól függ. Másrészt célszerűbbnek tartottuk egyelőre egy rendszer vizsgálatát, mivel az alapelvek és minták meghatározásához nagyon fontos a rendszeresség megállapítása.

A szövegek javítását 18 különböző személy végezte, mindegyikük felsőfokú, de különböző végzettséggel. Ebben a vizsgálatban fontos a végzettség, hiszen azt szeretnénk, ha a szövegeket minél jobb minőségben javítanák ki, tehát a szöveg lehetőleg ne tartalmazzon újabb nyelvi és helyesírási hibákat. A javításhoz útmutatásként azt kapták, hogy annyi javítást végezzenek csak el, amennyit feltétlenül szükségesnek érznek ahhoz, hogy a szöveg érthető legyen, de a szövegeknek nem kell feltétlenül helyesnek lenniük. Szándékosan nem kívántuk meghatározni az egyes kijavítandó szerkezeteket, mert arra is kíváncsiak voltunk, hogy az olvasók mikor tartják már éppen elfogadhatónak a szöveget. Ez az útmutatás nem vezet feltétlenül egységes javításhoz, mert az egyes javítók különböző mértékben érznek „kényszert” a szerintük nem helyes részek kijavításához, de ebből is levonhatunk tanulságokat, valamint a minimális utószerkesztés meghatározásához és az elfogadhatósághoz szükséges javítások megállapításához is segítséget nyújt. Sok esetben ugyanis többet javítanak ki, mint amennyi feltétlenül szükséges lenne az érthetőséghez (pl. ragok), de számukra a szöveg így lesz elfogadható. Ebből is azt vehetjük észre, hogy míg az érthetőség és az elfogadhatóság nem azonos kategóriák, sőt, egymásnak nem is szükséges és elégséges feltételei, az olvasók számára mégis összemosódnak.

A szövegek elérhetőségéhez készítettünk egy honlapot, ahonnan a javítók letölthették a szöveget. A program véletlenszerűen választ egy szöveget az adatbázisból, amelyet még nem javítottak ki, és azt ajánlja fel mentésre. Ha az olvasónak nem tetszik a szöveg témája, kérhet új szöveget, de nem tudja, hogy még milyen szövegek vannak. Így tehát a szövegek részben választhatóak. A honlap segítségével bárki letölthet szöveget, és vissza is küldheti, és minden javító annyi szöveget javít, amennyit szeretne. A javítók természetesen használhatták az



angol eredeti szövegeket, de kifejezetten kértük, hogy ne fordítsák újra a szöveget, az eredetit csak ellenőrzésre használják, valamint akkor, ha nem érthető mondattal találkozta, de akkor is csak a magyar fordítás elolvasása után.

A beküldött szövegeket ezután ellenőriztük, hogy valóban érthetőek-e, ellenőriztük, hogy mindent jól értelmeztek-e a javítók a szövegekben, tehát a szövegek pontosságát, valamint azt, hogy nem maradtak-e véletlenül nem kijavított részek. Az ellenőrzés után a magyar-magyar szövegeket a MemoQ illesztő moduljával illesztettük, és így párhuzamos szövegeket, azaz tmx kiterjesztésű fordítómemóriákat hoztunk létre. A fordítómemóriában a forrásnyelvi szegmensnek az eredeti, a célnyelvi szegmensnek pedig az utószerkesztett szöveg egy-egy mondata felel meg. Az így létrejött párhuzamos szövegeket a saját készítésű programmal dolgoztuk fel.

### **7.3.1.2. Az automatikus kivonatoló program**

Az utószerkesztési műveletek kiszűréséhez egy grafikus felhasználói felülettel rendelkező C# nyelvű alkalmazást készítettünk. A program bemenete egy párhuzamos fordítómemória. Ez a fordítómemória tetszőleges hosszúságú, tehát minden szöveghez tartozhat egy-egy fordítómemória, de tetszés szerint egy fordítómemória tartalmazhat több szöveget is. Mi az áttekinthetőség kedvéért az első módszert választottuk, tehát minden szöveghez létrehoztunk egy fordítómemóriát. A program beolvassa a dokumentumot, elkülöníti az eredeti és a javított szegmenseket, majd mindkettőt reguláris kifejezések segítségével szavakra, helyesebben tokenekre bontja. Minden szó, minden írásjel, valamint minden szám egy-egy token. A kötőjel kivétel az írásjelek alól, mivel nem számít egy különálló tokennek. Miután megvannak a tokenek, a program a Kis által javasolt módosított Levenstejn algoritmussal kiszűri a javítási műveleteket, azaz a beszúrt, a törölt, az áthelyezett és az átfirt tokeneket, valamint ezek pozícióját (Kis 2008: 78-81).

A Levenstejn algoritmus eredetileg két karakterlánc Levenstejn távolságát vagy szerkesztési távolságát adja meg. A szerkesztési távolság a szerkesztési műveletek száma, amelyre szükség van ahhoz, hogy az egyik karakterláncot úgy transzformáljuk, hogy pontosan a másikat kapjuk. Szerkesztési művelet a beszúrás, törlés és a helyettesítés (Navarro 2001: 35). Ebben az eredeti algoritmusban minden művelet karakterszintű. Ennek az algoritmusnak a módosított változatát írja le Kis, amelyben a két karakterlánc két mondat (azaz szegmens), a szerkesztési műveletek pedig token szintűek. Az algoritmus részletes és szemléletes bemutatását megtaláljuk Kisnél (Kis 2008: 78-82). A megvalósított algoritmus alapján a szerkesztési műveletek a következőket jelentik: A beszúrás azt jelenti, hogy a javított szegmensben az adott pozíción megjelenik egy új token, a törlés azt jelenti, hogy a javított szegmensben az adott pozíción nem található az adott token. A szegmens hossza egyenlő a benne található tokenek számával, a pozíció pedig a szegmensben belül a token sorszámát jelenti. A helyettesítés, azaz adott esetben az átírás azt jelenti, hogy az adott pozíción történt egy törlés és egy beszúrás is, az áthelyezés pedig azt jelenti, hogy a mindkét szegmensben megtalálható token nem ugyanazon a pozíción van. Az áthelyezésnél számított pozíció nem abszolút, tehát nem azt tekintjük áthelyezésnek, ahol a két egyező token nem azonos sorszámú a két szegmensben, hanem csak azt a helyzetet, amikor egy token a szegmens többi, „fő” részéhez képest megváltozott pozíción van, például a szegmens elejéről a végére kerül. A szegmens fő része az algoritmus szerint az a token vagy tokensorozat, amely az első illeszkedő pozíción van, vagy az első illeszkedő pozíciótól kezdődik.

Az algoritmus hátránya, hogy az átírásokat és egyben áthelyezéseket nem tudja kiszűrni, pedig erre nagyon sok példa fordul elő a szövegekben. Ezt nem is várhatjuk az algoritmustól, hiszen ez csak formai elemzéssel semmiképp nem oldható meg. Ha egy szót kicserélünk, és át is helyezzük más pozícióba, akkor az egyezést nem lehet megállapítani.

Az előbbi műveletekre nézzünk egy példát. A táblázat első sorában a javított szegmens, az első oszlopában pedig az eredeti szegmens van. A programban megvalósított algoritmus megállapítja az egyezéseket a két szegmensben, és azt is megállapítja, hogy csak részlegesen vagy teljesen egyeznek. Részleges egyezés esetén az illeszkedéseket tartalmazó ritka mátrixban azt is tárolja, hogy hány karakter volt megegyező. A program egyezésnek veszi azt, ha a tokenekben levő egyezés legalább a rövidebb token 50 százaléka, vagy rövidebb tokeneknél a két tokenben legalább 3 db karakter megegyezik, és a tokenek hossza között legfeljebb 5 karakter különbség van. Nézzünk meg egy mondatpárt:

Az eredeti mondat:

(9a) A vizsgálat azt találta, hogy a p21 gén cellákban volt következetesen tétlen a MRL egérfülből. (GF\_046)

Az utószerkesztett mondat:

(9b) A vizsgálat azt találta, hogy a p21 gén következetesen tétlen volt a sejtekben az MRL egér fülében. (USZ\_046)

(Az eredeti mondat tokenjei függőlegesen, az utószerkesztett mondat tokenjei pedig vízszintesen vannak a táblázatban.)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
		A	vizsgálat	azt	találta	,	hogya	a	p21	gén	következetesen	tétlen	volt	a	sejtekb	az	MRL	egér	fülb	.
1	A	1						1						1		1				
2	vizsgálat		1																	
3	azt			1																
4	találta				1															
5	,					1														
6	hogya						1													
7	a	1						1						1		1				
8	p21								1											
9	gén									1										
10	cellákban																			
11	volt												1							
12	következetesen										1									
13	tétlen											1								
14	a	1						1						1		1				
15	MRL																1			
16	egérfülből																			
17	.																			1

5. ábra: Szerkesztési műveletek az eredeti és utószerkesztett mondatban

A példamondatban a beszúrt tokenek a következők: *sejtekb*, *az*, *egér*, *fülb* (kék háttér), a törölt tokenek: *cellákban*, *egérfülből* (sárga háttér), az áthelyezett tokenek: *következetesen*, *tétlen* (zöld háttér). Olyan, azonos pozíció levő átírás nincs a két mondatban, amelyet az algoritmus ki tudna szűni. A *cellákban* → *sejtekb* átírás lenne, de nem azonos pozíciókon, az *egérfülből* → *egér fülb* pedig a pozíció mellett a tokenek számában sem egyezik. A program mégis segít az átírások kiszűrésében, mégpedig úgy, hogy ha a törléseket és a beszúrásokat egymás mellé tesszük mondatpáronként, és környezetükkel együtt vizsgáljuk őket, akkor kézi vizsgálattal gyorsabban megállapíthatóak az átírások, mint ha a mondatokat kellene átvizsgálni.

Mivel egy szegmensben belül többször is előfordulhat ugyanaz a token, valamint mivel a részleges egyezéseket is tároljuk, könnyen előfordul, hogy egy soron vagy oszlopon belül több egyezés is van. A mátrixban is láthatunk erre példát: 3 határozott névelő is van az eredeti mondatban, ezek mindegyike

illeszkedik minden határozott névelővel az utószekesztett mondatban. Ezek közül a szerkesztési műveletek megállapításához mindenképpen csak egy maradhat, hiszen a szegmensekben egy tokennek csak egy párja lehet. A megvalósított algoritmus a következőképpen állapítja meg, hogy melyik egyezés maradhat meg: Az első szempont az, hogy a legjobb egyezés maradjon meg, tehát kiválasztjuk a maximum értéket egy soron belül, és ezt tartjuk meg egyezésként, az összes többi érték 0 lesz. Ha több egyforma érték van, akkor azt az értéket tartjuk meg, amelyiknél a két szegmensben a pozíciók a legközelebb esnek egymáshoz, tehát a sor és az oszlop sorszáma között a legkisebb a különbség. Amennyiben ez is megegyezik, akkor az első értéket tartjuk meg. Ezt a műveletet megismételjük minden sorra, de ha a maximum olyan oszlopban van, ahol már van érték, akkor vagy másik maximumot választunk a sorból, ha pedig nincs ilyen, akkor úgy vesszük, mintha nem lenne egyezés. A példamondatban a törölt egyezéseket narancssárgával jelöltük, és lilával azokat, amelyeket megtartottunk.

Ha a mátrix elkészült, ebből a Kis által javasolt algoritmus segítségével kiszűrjük a javítási műveleteket. A program minden adatot adatbázisban tárol, ahonnan nagyon egyszerűen exportálható a kívánt adathalmaz. Az adatok jobban áttekinthető megjelenítéséhez SQL lekérdezések készültek.

### **7.3.1.3. Általános elemzés**

A szövegekben szereplő tokeneken szófajelemzést végeztünk. A szófaji elemzést részben kézzel végeztük, részben pedig az adatbázis segítségével automatikusan. A kategorizáláshoz használt szófajok: főnév, tulajdonnév, ige, melléknév, főnévből képzett melléknév, számnév (határozott, határozatlan és évszám), melléknévi igenév, határozószó, főnévi igenév, határozói igenév, névmás (típusok alapján, ezenkívül utalószerben megkülönböztetve), módosítószó, kötőszó, névutó, igeekötő, névelő (határozott és határozatlan) írásjel, angol eredeti, mértékegység.

Elsőként azonnal látjuk, hogy az utószerkesztett szövegekben a tokenek száma (26134) alig tér el az eredeti fordítások tokenszámától (26253), amely csak 0,5 százalékos negatív irányú eltérés, tulajdonképpen elhanyagolható.

A szófajok megoszlása az eredeti és az utószerkesztett szövegekben a következő:

Szófaj	Eredeti	Utószerkesztett
fn	4907	5058
ír	3827	4009
ige	3269	3088
nvh	2357	2538
nm	2027	1871
mn	1898	1879
fn, tul	1820	1678
köt	1785	1826
szn	946	903
nvl	626	494
hat	525	605
mód	441	475
nu	418	429
in fn	322	286
ik	292	275
fn/mn	229	281
in mn	192	180
angol	141	0
mért	131	134
?	57	17
in hat	25	81
röv	12	12
rag	6	12

**22. táblázat: Szófajok megoszlása a két korpuszban**

Az eredetiben angolul szereplő szavak száma nullára csökkent, ahogy ez el is várható. A nem bekegategorizálható, általában érdekes szóösszetételek, egybeírt szavak száma jelentősen csökkent, 70%-kal. Talán azt várnánk ezektől, hogy számuk 0-ra csökkenjen, de a javítók sokszor úgy vélhették, hogy ezek az érthetőséget nem zavarják. A főnevek száma az utószerkesztett szövegekben nőtt, 3%-kal, az igék száma viszont 5,5%csökkent. Arányaiban a legnagyobb változás a

határozói igeneveknél van, amelyek száma több mint kétszeresére nőtt. Nézzünk erre példát:

(10a) Még korai napok van, és Londonnak van egy olyan éghajlata, ami kicsit közelebb azokhoz a feltételekre, amikhez a redstartsok\* **hozzászoknak** a kontinensen. (GF\_066)

(10b) Még korán van, és Londonnak van egy olyan éghajlata, ami kicsit közelebb van azokhoz a feltételekre, amikhez a rozsdafarkúak **hozzá vannak szokva** a kontinensen. (USZ\_066)

A melléknévek száma szinte ugyanannyi maradt, egészen elhanyagolható mértékben csökkent, a főnévből képzett melléknévek száma viszont viszonylag jelentősen, 23%-kal nőtt.

(11a) Ott a félelmek azok az egyetemekre gyakorolt kormányzati nyomások voltak hogy többet vegyenek, a munkásosztályból való diákok jó-képzett **magániskola-tanulóktól** tagadnának meg helyeket Oxfordnál és Cambridge-nél. (GF\_047)

(11b) Félelmek vannak, hogy az egyetemekre gyakorolt kormányzati nyomás miatt, hogy többet vegyenek fel a munkásosztályból való diákokból, jól képzett **magániskolai** tanulóktól tagadnának meg helyeket Oxfordban és Cambridge-ben. (USZ\_047)

A határozatlan névelők száma 21%-kal csökkent, míg a határozott névelők száma nőtt, bár nem ennyire nagy mértékben, 7,7%-kal. A határozatlan névelők csökkenése nem meglepő, ha figyelembe vesszük az eredeti forrásnyelv, az angol és a magyar eltérő névelőhasználati jellegzetességeit.

A tulajdonnevek száma 7,8%-kal csökkent, amely meglepő is lehetne, de ennek oka az, hogy az intézménynevek, egyesületek, társaságok, alapítványok nevei a nyers gépi fordításokban nincsenek lefordítva, de az érthetőség kedvéért sokat ezek közül mégis célszerű lefordítani. A főnévi és melléknévi igenevek száma csökkent, a főnévi igenevek nagyobb mértékben, 11%-kal, a melléknévi igeneveké 6%-kal. A melléknévi igenevek számából nem vonhatunk le következtetéseket, mivel a melléknévi igenév melléknévi szerepet tölt be, az elemzés során sokszor a melléknév kategóriába került.

Há a névmásokat tovább kategorizáljuk, akkor láthatjuk, hogy a birtokos névmások száma csökkent arányaiban a legnagyobb mértékben, 100%-ban,

ugyanis az utószérkesztett szövegekben egyetlen birtokos névmás sincs. Ennek oka, hogy az eredetiben a birtokos névmás ilyen helyzetekben fordul elő:

(12a) Heath Ledger, 28, aki januárban meghalt egy véletlen receptgyógyszertúladagolásban, elhagyott mindent az **övé** 30 millió dolláros **szerencsét** becsült a szüleinek és nővéreinek egy írott azelőtt elkezdte keltezni Michelle Williams színésznőt. (GF\_007)

(12b) Heath Ledger, 28, aki januárban meghalt egy véletlen gyógyszer-túladagolásban, itt hagyott mindent a 30 millió dollárra becsült **vagyont** a szüleinek és nővéreinek egy végrendeletben, mielőtt elkezdett járni Michelle Williams színésznővel. (USZ\_007)

A kérdő névmások száma lecsökkent 35,5%-kal, a vonatkozó névmások és a kötőszavak száma ezzel párhuzamosan viszont nőtt. A hibakategóriák felállítása után ez az eredmény nem meglepő, hiszen az érthetőség szempontjából az egyik legsúlyosabb hibának az bizonyult, amikor a vonatkozó névmás helyett a mondatban kérdő névmás szerepelt. A kérdő névmás ugyan szerepelhet vonatkozói szerepben magyar szövegekben, azaz lehet vonatkozó névmás is, de ha emellett a szórend kérdő mondatra, tagmondatra utal, ez félrevezeti az olvasót.

(13a) Ez esőt fog hozni csütörtökön **melyik** lesz nehéz délnyugaton Anglia és Walesben. (GF\_057)

(13b) Ez esőt fog hozni csütörtökön, **amelyik** heves lesz Anglia délnyugati részén és Walesben. (USZ\_057\_2)

A személyes névmások száma szintén jelentősen csökkent, 27%-kal, amiért az angol és a magyar közti különbség okolható. A névmások részletes megoszlása a következő táblázatban látható.



A névmások részletes felosztása:

Eredeti		Utószerkesztett	
nm	44	nm	57
nm, ált	49	nm, ált	45
nm, birt	11	nm, birt	0
nm, hatlan	77	nm, hatlan	70
nm, kérd	110	nm, kérd	71
nm, kölcs	4	nm, kölcs	6
nm, mut	341	nm, mut	278
nm, szem	341	nm, szem	249
nm, utal	428	nm, utal	442
nm, vissza	41	nm, vissza	53
nm, von	581	nm, von	600

23. táblázat: A névmások megoszlása

#### 7.3.1.4. Utószerkesztési műveletek szerinti elemzés: beszúrások és törlések

Az eredetiből törölt és beszúrt szavak szófaji megoszlása (az átírt szavak nélkül) így alakult. Az alábbi táblázat minden olyan tokenet tartalmaz, amely az eredeti szövegben nincs benne. Ez azt is jelentheti, hogy a javított szövegben nem jelent meg más a token helyén, de azt is, hogy ez egy átírt token. Részben ki lehet szűrni ezeket az átírt tokeneket úgy, hogy megnézzük a törölt szó pozícióját, majd azt, hogy a beszúrások között ugyanazon a mondatpozíción történt-e beszúrás. Ha igen, akkor az ilyen tokenek nagy valószínűséggel az átírt tokenek, de ez a módszer nem teljesen biztos, hiszen a javított szövegben ugyanazon a pozíción már egészen más (szerepű) token is szerepelhet az áthelyezések, előző törlések, beszúrások miatt. Ha az azonos pozíción nem volt beszúrás az utószerkesztett szövegben, akkor ezt „teljes” törlésnek hívjuk. A másik oldalról a beszúrt tokenekkel ugyanez a helyzet, itt a „teljesen” beszúrt szó pedig olyan token, amely az utószerkesztett szövegben megjelent úgy, hogy nem történt azonos pozíció átírás, tehát az eredeti szövegben semmilyen formában nem szerepelt. Az algoritmus sajnos nem „veszi észre” az áthelyezett és egyben átírt tokeneket, tehát

az ilyenek is szerepelnek a törlések és a beszúrások között. Mivel a teljes törlések és beszúrások kiszűrése nem egészen megbízható, nézzük meg a teljes és nem teljes törléseket illetve beszúrásokat. A továbbiakban viszont a törlések és beszúrások vizsgálatakor a tágabb értelmükben vesszük ezeket a műveleteket, és az elemzésben is már csak ezek szerepelnek. Először mindig a darabszámok szerepelnek, a következő oszlopban pedig a törölt tokenek aránya az eredeti tokenekhez képest, valamint a beszúrt tokenek aránya az utószerkesztett tokenekhez képest.

Szófaj	Törlések		Teljes törlések		Beszúrások		Teljes beszúrások	
	Db	%	Db	%	Db	%	Db	%
?	33	57,89%	17	29,82%	8	47,06%	2	11,76%
angol	119	84,40%	35	24,82%	0	0,00%	0	0,00%
fn	933	19,01%	476	9,70%	1029	20,34%	511	10,10%
fn, tul	263	14,45%	108	5,93%	169	10,07%	81	4,83%
fn/mn	25	10,92%	12	5,24%	47	16,73%	21	7,47%
hat	119	25,27%	68	14,44%	119	21,56%	70	12,68%
ige	904	27,65%	465	14,22%	758	24,55%	354	11,46%
ik	128	43,84%	72	24,66%	112	40,73%	70	25,45%
in fn	111	34,47%	46	14,29%	81	28,32%	38	13,29%
in hat	8	32,00%	6	24,00%	41	50,62%	20	24,69%
in mn	60	31,25%	38	19,79%	51	28,33%	32	17,78%
ind	1	100,00%	0	0,00%	0	0,00%	0	0,00%
ír	410	10,71%	289	7,55%	604	15,07%	408	10,18%
köt	236	13,22%	152	8,52%	271	14,84%	157	8,60%
mért	46	35,11%	23	17,56%	49	36,57%	27	20,15%
mn	344	18,12%	154	8,11%	397	21,13%	182	9,69%
mód	55	12,50%	28	6,36%	96	20,21%	52	10,95%
nm	3	6,82%	2	4,55%	13	22,81%	4	7,02%
nm, ált	12	24,49%	9	18,37%	6	13,33%	4	8,89%
nm, birt	11	100,00%	7	63,64%	0	0,00%	0	0,00%
nm, hatlan	18	23,38%	10	12,99%	12	17,14%	3	4,29%
nm, kérd	41	37,27%	22	20,00%	13	18,31%	7	9,86%
nm, kölcs	0	0,00%	0	0,00%	2	33,33%	1	16,67%
nm, mut	124	36,36%	71	20,82%	77	27,70%	56	20,14%
nm, szem	122	35,78%	71	20,82%	32	12,85%	17	6,83%
nm, utal	102	23,83%	51	11,92%	123	27,83%	68	15,38%

Szófaj	Törlések		Teljes törlések		Beszúrások		Teljes beszúrások	
	Db	%	Db	%	Db	%	Db	%
nm, vissza	5	12,20%	1	2,44%	15	28,30%	12	22,64%
nm, von	143	24,61%	82	14,11%	167	27,83%	91	15,17%
nu	85	21,79%	54	13,85%	104	26,53%	66	16,84%
nu - mn	7	25,00%	4	14,29%	14	35,90%	6	15,38%
nvh	333	14,13%	233	9,89%	491	19,35%	298	11,74%
nvl	142	22,68%	99	15,81%	21	4,25%	9	1,82%
rag	3	50,00%	1	16,67%	8	66,67%	5	41,67%
röv	0	0,00%	0	0,00%	0	0,00%	0	0,00%
szn, hat	109	14,34%	47	6,18%	84	11,46%	38	5,18%
szn, hatlan	32	17,78%	20	11,11%	19	11,66%	10	6,13%

**24. táblázat A törölt és a beszúrt tokenek**

Nézzünk meg néhány érdekességet a táblázat alapján. Az előző táblázat után nem meglepő módon a birtokos névmásokat törölték, és egyetlen egy sem volt beszúrva, és ugyanez a helyzet az eredetileg angol szavakkal is. A nem meghatározható szavak (kérdőjeles) beszúrára a magyarázat az, hogy az utószerkesztők véletlenül egybe írtak szavakat, például *támogatjaés, mivelta, mégiseléggé*. Indulatszó egyetlen egy volt az eredeti szövegekben: *nahát*, és azt is kitörölték. Nagy arányban töröltek ki ragokat, de ugyanígy be is szúrták őket, de ezeknek a száma nagyon alacsony. Ragnak hívtuk az olyan tokeneket, amelyek kötőjellel kapcsolódtak valamilyen elemhez, általában mértékegységhez – mivel a kötőjel a tokenekre bontáskor egyébként nem számít külön tokennek, csak akkor választ szét két tokent, ha közben van egy nem betű karakter – például *90%-ával*. A személyes és kérdő névmások több mint 30%-át törölték, de csak 12 illetve 18 százalékot szúrtak be. A mutató névmásoknak is több mint 30%-a törölt, de cserébe szintén majdnem 30% be is van szúrva. Ezeknek a száma is elég nagy, 124 illetve 77. Ez azt jelenti, hogy az eredeti szövegekben a mutató névmásokra figyelni kell.

Szinte minden törölt mértékegység helyett meg is jelent javított mértékegység az utószerkesztésben, és a javított mértékegységek aránya is elég nagy: a

szövegben szereplő mértékegység 35%-a. Megjegyzésként szeretnénk hozzáfűzni, hogy saját szubjektív meglátásunk szerint ennél még több mértékegység is javításra szorulna a könnyebb érthetőség érdekében. Az eredetileg angolul szereplő mértékegységeket ugyan legtöbbször kijavították az utószerkesztők, de az angol rövidítések közül néhány megmaradt, pl. *lb*, *st* vagy akár *stone*. Ezeket feltétlenül ki kellene javítani, és magyarul teljesen kiírni. Mivel ezek reáliák, az átváltásukkal kapcsolatban lehetnek különböző álláspontok. Ha szakszövegek fordításáról van szó, akkor kötelező a pontos átváltás, azaz a teljes átalakítás, de műfordítások esetében az átváltással elveszhet a kulturális színezet (Tellingner 2007: 30, Klaudy 2005: 134). Már az elején leszögeztük, hogy a gépi fordítás műfordításra egyáltalán nem alkalmas, még a minimálisan, sőt, véleményünk szerint még a teljesen utószerkesztett változata sem, tehát itt nem lehet kérdés a teljes átalakítás. A jelenleg használt korpusz viszont egyrészt jellegzetesen a brit kultúrából származó cikkeket tartalmaz, azaz lehet, hogy az olvasó várná az eredeti mértékegységeket. Másrészt pedig a nem átalakított mértékegységeket az olvasó is át tudja számolni, ha az információt fontosnak tartja. Tehát a mértékegységek átváltását nem tartjuk elengedhetetlennek, de mégis javasoljuk. Ennek több indoka is van: egyrészt nem garantált, hogy az utószerkesztett szöveg típusa hasonló lesz a későbbiekben, másrészt a feldolgozási erőfeszítést jelentősen csökkenthetjük, ha nem az olvasónak kell elvégeznie az átváltást. Harmadrészt hiába ismertek a mértékegységek nevei magyarul is, az olvasó gyakran mégsem tud konkrét értéket kapcsolni hozzájuk.

A mértékegységeknél még nagyobb arányban változtak az igekötők. A törölt igekötők aránya 44%, a beszúrtaké pedig 41%. Főnévi és melléknévi igeneveket is nagy arányban töröltek és szúrtak is be, 30% körül. A törölt igék és a főnevek száma szinte megegyezik, de tekintettel arra, hogy a főnevek száma nagyobb, a törölt igék aránya 27%, míg a törölt főnevek az eredeti főneveknek csak 19%-át teszik ki. Ezzel szemben viszont több beszúrt főnév van, mint törölt, igéből viszont kevesebb van, így az arányuk a beszúrásokban közelít egymáshoz: 20% és 25%.

Nagy a különbség a törölt és a beszúrt határozatlan névelők között: 23% és 4%, ezt láttuk a szófajok első vizsgálatakor is. A határozott névelők arányában nincs ilyen nagy különbség, de a különbség éppen fordított: a beszúrt határozott névelők száma és aránya nagyobb a törölteknél (19% és 14%). Az írásjelek aránya a teljes szövegben nagy, a második leggyakoribbak a főnevek után, javítási arányuk mégis nagyon kicsi, 11%-ot töröltek, és 14%-ot szúrtak be. Ez azt mutatja, hogy nem az írásjelek jelentik a legnagyobb problémát. Az írásjeleken belül az idézőjeleket érdemes kiemelni, ezeknek a száma változott a legjobban, méghozzá úgy, hogy 4-ről 104-re változott a számuk. Az eredeti szövegekben ugyanis a párbeszéd, idézetek nincsenek jelölve, így az utószerkesztők ezeket sokszor idézőjelekkel jelölték az utószerkesztésben. A vesszők száma összességében a legmagasabb, de ezek száma nem változott jelentősen, csak 4%-kal, de a törlések aránya 14%, a beszúrt vesszők aránya pedig 18%, és bár az áthelyezések elemzése ezután következik, érdemes itt kiemelni, hogy az áthelyezett vesszők aránya 28%. Ezek alapján érdemes a vesszőket javítani, mivel utószerkesztésük egyszerű, de a feldolgozási erőfeszítést ehhez képest nagyban lecsökkentik a megfelelően elhelyezett vesszők. A többi írásjel változása nem jelentős.

A következő táblázat legnagyobb számban törölt, beszúrt 25 tokent, valamint a legnagyobb arányban törölt és beszúrt 25 tokent mutatja. Az arány azt mutatja meg, hogy ahhoz képest, hogy abból a tokenből összesen mennyi volt, mennyit töröltek és szúrtak be.

Legnagyobb számban törölt		Legnagyobb arányban törölt		Legnagyobb számban beszúrt		Legnagyobb arányban beszúrt	
,	315	Barrier	13	,	332	'	12
a	219	Great	13	a	263	férfiak	9
egy	145	Vándorárus	11	az	141	Korall	9
az	118	övék	7	.	137	lehetővé	8
hogy	103	hosszában	5	egy	111	férfi	6
*	58	Igazságos	5	hogy	109	amelyek	5
azt	56	ütni	5	és	87	férfiaknál	5
miközben	48	ütötte	5	azt	42	munkaügyi	5

volt	48	alakok	4	nem	39	során	5
ez	46	enyém	4	volt	30	adtak	4
ki	37	erő	4	ez	27	anyukája	4
meg	36	gazdálkodásnak	4	*	25	eset	4
van	31	irányító	4	miközben	24	fogadnia	4
amit	31	írott	4	ki	22	fonttal	4
és	28	kedvel	4	Mondta	21	képes	4
el	24	megengedte	4	ahogy	19	láb	4
által	23	megy	4	de	18	mezőgazdaságnak	4
levő	23	súly	4	meg	17	miszerint	4
Olyan	21	abból	3	vagy	16	oda	4
voltak	20	cellák	3	van	16	terményeket	4
£	19	csomót	3	én	14	Védelmi	4
annak	19	ejtsék	3	-	14	adta	3
nem	18	elkövetni	3	több	13	alól	3
ami	17	elülső	3	olyan	13	amelyben	3
ő	16	embereknél	3	amit	12	bomlanak	3

**25. táblázat: A leggyakrabban törölt és beszúrt tokenek**

Érdekes megnézni, hogy a számban leggyakrabban törölt és beszúrt szavak nem az alapszófajok közül kerülnek ki, valamint a két lista nagyon hasonló. Azok a szavak viszont, amelyeket arányaikban leggyakrabban változtattak, néhány kivétellel alapszófajú szavak, és ez a két lista egyáltalán nem hasonlít egymásra.

Az eredmények többet mondanak, ha megnézzük a tokenek környezetét. Kiszűrtük a fent említett leggyakrabban törölt és beszúrt tokenek közül néhányat a 3 szavas környezetét. Ezt megvizsgálva láthatjuk, hogy a törölt és beszúrt lista sok olyan szót tartalmaz, amelyek egymásnak javításai. Mivel a lista nagyon hosszú, csak néhány példát mutatunk be. A kontextusban tokenek között szóköz van.

Törölt szó	Eredeti környezet	Javított környezet
embereknél	alatt több az embereknél . Ők különösen	éjszaka alatt a férfiaknál különösen Ők .
embereknél	húsz perc az embereknél alszik jó éjt	perccel több alvásra van férfiaknál a szükségük
embereknél	, hogy az embereknél keményebben működtessék az	, hogy a férfiaknál az működtessék keményebben
alakok	lesz , az alakok javasolják azt	lesz , a számok azt mutatják ez

Törölt szó	Eredeti környezet	Javított környezet
	valamennyiért	
alakok	mondott : Az alakok , amiket használtak	mondta : Az adatok , amiket használtak
alakok	, hogy az alakok kihangsúlyozták az igényt	azt az igényt , hogy törvényes jogokat
alakok	vannak . Az alakok mutatják a potenciálisan	vannak . A számok mutatják a potenciálisan
cellák	hogy a felnőtt cellák mint embrionális őssejtek	, hogy a felnőtt sejtek mint embrionális
cellák	viselkedjenek - a cellák egy pluripotential rendelkező	őssejtek viselkedjenek – az embriókban egy pluripotens
cellák	A nagyon megújító cellák egy növekedésének a	Az erősen megújuló sejtek növekedésének és az
elkövetni	A prioritásuk van elkövetni a jövőig kapsold	A prioritás : eldönteni a jövőben a
elkövetni	hogy kit akarnak elkövetni , pénzügyi stabilitás	a pénzügyi stabilitás és saját otthon .
elkövetni	vagy az érzelmekben elkövetni túl sokat ,	magát időben vagy érzelmekben , amelyikek becsesek
övék	voltak osztva , övék több millió Tesco	, amiket a Tesco osztott ki ,
övék	voltak osztva , övék több millió Tesco	lettek kiosztva a Tesco által , károsabbak
övék	kezdődött . Nép övék 80 több nagylelkű	kezdődött . A 80-as éveiben levő emberek
övék	összeadnak : Gyakorló övék a választás szabadsága	hozzáteszi : Abban , hogy a választás
Övék	a laborban . Övék több gyűjteménye 1	a laborban . A több mint 1
övék	egy telefonmérnök felfedezett övék biztonságba helyez nyolc	meg , mikor egy telefonmérnök felfedezte a
övék	kell beházasodniuk , övék húsz vagy szembenéz	a nőknek 20-as éveikben meg kell házasodniuk
hosszában	történni . Mexikó hosszában levő Kate Winslet	történni . Kate Winslet Mexikó felé szökik
hosszában	£ 900 adósság hosszában levő FarmVille felhasználói	FarmVille felhasználó 900 adósságot csinált . Egy
hosszában	csapadék és tömítés hosszában levő áztatás árad	felszívásában és az áradás megakadályozásában ] ,
hosszában	voltam négy krémsütemény	ettem négy krémsüteményt

Törölt szó	Eredeti környezet	Javított környezet
	hosszában levő noshing *	együltömben , mondván ,
hosszában	Négy a sorozat hosszában levő év mi	Négy egymást követő év és mi mindketten
igazságos	vegyen , nem igazságos egy ingerlékeny lány	vegyen , nem csak egy bátor lányt
igazságos	volt az alkalom igazságos , hogy van	soha nem volt olyan alkalom , hogy
igazságos	most . Én igazságos vagyok miközben gyakorlati	. Én csak gyakorlatias vagyok , hogy
Igazságos	? Spenót . Igazságos spenót . Születésnap	ebédre spenót . Csak spenót . Születésnapos
igazságos	üzenet népnek van igazságos , hogy gondos	üzenet az embereknek az , hogy óvatosság
ütötte	. A haditengerészetet ütötte egy sor botrány	. A haditengerészet egy sor botrányt élt
ütötte	a házasságot hogyan ütötte egy tornádó egy	, “ mint egy tornado ” mikor
ütötte	a Vauxhall Vectra ütötte a vizet .	a Vauxhall Vectra elérte a vizet .
ütötte	egy mindenható csattanással ütötte a vizet és	egy mindenható csattanással csapódott a vízbe és
ütötte	a pár kapcsolata ütötte a sziklákat ,	kitudódott , hogy a pár kapcsolata szikláknak

**26. táblázat: A leggyakoribb törölt szavak környezete**

Nézzük a beszúrt szavakat:

Beszúrt szó	Javított környezet	Eredeti környezet
férfi	volt és a férfi tudta pontosan ,	emberi van ő ? Ő 17 éves
férfi	mondta . Új férfi rám nézett rémülettől	Új ember , akire néztek , én
férfi	az országokból fotópárokat férfi arcokkal . Mindegyik	az országokból a férfiarok több pár fotóját
férfi	A 45 éves férfi , aki az	Az 45 éves , ki az utolsó
férfi	most több nőtlen férfi van Ez az	most nőtlenebb emberek vannak Ez az egyedülálló
férfi	. Több egyedülálló férfi van Angliában és	. Több egyedülálló ember van Angliában és
férfiak	. A nőtlen férfiak a 75 év	. Az egyedülálló emberek a hivatalnok szerint
férfiak	síránkozása : a férfiak közül ,	síránkozása : az emberek



Beszúrt szó	Javított környezet	Eredeti környezet
	akivel	közül annyira mindegyik
férfiak	, azaz a férfiak és nők különböző	végzett tanulmány , az emberek és nők
férfiak	. Másfelől a férfiak , akik eltévedve	. Az emberek , akik másfelől sodródva
férfiak	a férfias kinézetű férfiak vonzósága fogyatkozott ,	férfias-néző emberek vonzalma fogyatkozott , a folyóirat
férfiak	nők ízlése a férfiak terén lágyul .	női ízlése az emberek terén lágyul .
férfiak	nők ízlése a férfiak terén lágyul ,	ízlését az emberek terén annyira lágyul macsó
férfiak	funkcióját mint a férfiak és tovább alszanak	az emberek tartják meg az agykéreg funkcióját
férfiak	alvásukból mint a férfiak . Tehát sok	mély alvásként tartanak meg az alvásukból többet
férfiaknál	, hogy a férfiaknál keményebben működtessék az	, hogy az embereknél keményebben működtessék az
férfiaknál	éjszaka alatt a férfiaknál . Ők különösen	alatt több az embereknél . Ők különösen
férfiaknál	van szükségük a férfiaknál de sok nem	embereknél alszik jó éjt ! de sok
férfiaknál	átlagos házasulandó kor férfiaknál 36 . 5-re	átlagos kor házasságnál nőknek emelkedett 36 .
férfiaknál	az átlagos kor férfiaknál éppen meghaladja 32	az átlagos kor embereknek épp több mint
fogadnia	mezőgazdaságnak el kell fogadnia a GM technológiát	gazdálkodásnak GM technológiát kell megölelnie hogy 21
lehetővé	látványos kilátást tesz lehetővé a Föld és	kilátásait és teret fogja nyújtani . Ez
lehetővé	, a viszonyok lehetővé teszik neki ,	, az ügyek megengedik neki , hogy
lehetővé	emberekkel általában - lehetővé teszi nekem ,	rúgkapál általában - megengedi nekem , hogy
lehetővé	lehet , hogy lehetővé teszik ezeknek a	lehet engedd meg ezeknek a szervezeteknek a
lehetővé	sokat , ami lehetővé tette a magániskoláknak	sokat , ami megengedte a magániskoláknak ,
lehetővé	filmnél , ami lehetővé tette a	megengedte a kutatóknak ,

Beszúrt szó	Javított környezet	Eredeti környezet
	kutatóknak	hogy létrehozzanak egy
lehetővé	szándékok értelmezése ráadásul lehetővé tehetné a rendőrségnek	szándékok értelmezése ráadásul megengedhetné a rendőrségnek ,

**27. táblázat: A leggyakoribb beszúrt tokenek környezete**

Most pedig nézzük meg az áthelyezett tokenek szófaji megoszlását, és az arányukat az eredetiben levő számokhoz képest:

Szófaj	Darabszám	Arány
?	8	14,04%
angol	4	2,84%
fn	670	13,65%
fn, tul	304	16,70%
fn/mn	43	18,78%
hat	70	14,86%
ige	373	11,41%
ik	26	8,90%
in fn	37	11,49%
in hat	4	16,00%
in mn	25	13,02%
ind	0	0,00%
ír	724	18,92%
köt	288	16,13%
mért	29	22,14%
mn	285	15,02%
mód	54	12,27%
nm	7	15,91%
nm, ált	5	10,20%
nm, hatlan	9	0,00%
nm, kér	1	11,69%
nm, kérd	12	11,82%
nm, kölcs	1	25,00%
nm, mut	47	13,78%
nm, szem	32	9,38%
nm, utal	49	11,45%
nm, vissza	4	9,76%
nm, von	90	15,49%
nu	65	16,67%
nu - mn	3	10,71%

nvh	571	24,23%
nvl	138	22,04%
rag	0	0,00%
röv	0	0,00%
szn, hat	163	21,45%
szn, hatlan	31	17,22%

**28. táblázat: Az áthelyezett tokenek szófaji megoszlása**

A legnagyobb arányban a kölcsönös névmások kerültek új pozícióba, de mivel a számuk annyira alacsony, ebből semmilyen következtetést nem vonhatunk le. A második helyen álló szófaj, amelyet áthelyeztek, a határozott névelő, majd a mértékegységek, a határozatlan névelők, és a határozott számnevek. A mértékegységek és a számnevek valóban nagy gondot okoznak az utószerkesztésben, mert ezek gyakran nagyon messze kerülnek a hozzájuk tartozó elemektől, és nagyon nehéz kikövetkeztetni, mire vonatkoznak.

(14a) Britannia **74-nél** többen sütkérezett órák napsütés március 1-je és 14-e és körülbelül **77-e** között az egész havi átlag százaléka. (GF\_047)

(14b) Britannia sütkérezett a több, mint **74** napsütéses órában március 1-je és 14-e között, ez körülbelül az egész havi átlag **77** százaléka. (USZ\_047)

Nézzük meg a leggyakrabban áthelyezett 25 tokent. Az arányaiban leggyakrabban áthelyezett tokeneknél ez az arány mindenhol 100%, ezért ezt nem tüntettük fel.

Token	Darabszám	Token	Darabszám
,	638	megsebesítettek	3
a	430	52	2
az	163	65	2
egy	130	905	2
hogy	121	2F	2
és	119	apoptosis	2
.	38	átgondolnak	2
volt	33	baktériumok	2
nem	24	Chivers	2
Azt	22	csaló	2
-	19	erőforrások	2
vagy	16	esküvőben	2
több	15	évtized	2
mondta	14	Fayed	2
ez	13	használó	2

Token	Darabszám	Token	Darabszám
de	12	Helen	2
amit	10	hétfőn	2
éves	10	hőmérsékleteket	2
kell	10	hüvelykjét	2
két	10	igénnyel	2
mondták	9	játék	2
miután	9	jogokat	2
olyan	9	jómódú	2
amikor	9	készültek	2
voltak	9	kiáltanak	2

**29. táblázat: A leggyakrabban áthelyezett tokenek**

Az eredmény jellege nagyon hasonló, mint a törölt és beszűrt tokenek esetében. Az áthelyezési listán viszont előkelő helyet foglalnak el a számok, mivel – ahogy már korábban is említettük – a számok nagyon gyakran a szövegben ránézésre teljesen véletlenszerű pozíciókon helyezkednek el. A számok számjeggyel írva viszont mindig pontosak, tehát törölni, beszűrni nem kell őket. Nézzük meg néhány áthelyezett tokenek környezetét is:

Áthelyezett szó	Eredeti környezet	Javított környezet
52	a francia , 52 , amivel 1981-ben	, és French 52 akivel 1981-ben találkoztak
52	autóval fizető £ 52 egy hónap több	család két autóval 52 fonttal többet fizet
65	mi százalékom vagy 65 vagy öregebb ,	Tizenhat százalékunk vagy 65 vagy öregebb ,
65	nehéz tüzelőolajat és 65 , 0 tonna	nehéz tüzelőolajat és 65 , 0 tonna
905	fia költötte £ 905 FarmVille-en . Kiürítette	hogy a fia 905 fontot költött a
905	költ van £ 905 , de a	A teljes költség 905 font körül van
baktériumok	melyikeket győzik le baktériumok trágyakupacokban . Tegnap	a trágyában lévő baktériumok bontják le .
baktériumok	melyikeket győzik le baktériumok trágyakupacokban . Tegnap	a komposztban lévő baktériumoktól bomlanak le .

Áthelyezett szó	Eredeti környezet	Javított környezet
csaló	gyanún tartóztatták le csaló . West Midlands	25 snookeresét , család gyanújával tartóztatták le
csaló	11-én , csütörtökön csaló , mondta West	35 éves embert család gyanújával tartóztattak le
erőforrások	rendőrőrszobánál lesznek az erőforrások folytonosan monitor mi	rendőrőrszobánál lesz annyi erőforrás , hogy folytonosan
erőforrások	ahogy az összemelő erőforrások ösztönzik az energiavállalatokat	ahogy az összemelő erőforrások ösztönzik az energiavállalatokat
évtized	, csak egy évtized alatt három évig	először házasodik egy évtized alatt csaknem három
évtized	, mivel két évtized körül , képtelenek	mondták , két évtizede képtelenek voltak hordani
játék	legutóbbi Premier League játék bevonja Stoke-ot ,	legutóbbi Bajnokok Ligája játékot ahol Stoke játszott
játék	Zyngát tavaly a játék , és figyelmeztetett	alapítója , a játék pénzesítéséért kritizálta Zyngát
megsebesítettek	gyereket kritikusan tegnap megsebesítettek , amikor az	Két gyerek kritikusan megsebesült tegnap , amikor
megsebesítettek	84-et komolyan 2008-ban megsebesítettek . Zöld ,	és 84-en komolyan megsebesültek 2008-ban . A
megsebesítettek	az elmúlt évtizedben megsebesítettek az útjainkon ,	megöltek vagy komolyan megsebesítettek az útjainkon az

**30. táblázat: A leggyakrabban áthelyezett tokenek környezete**

#### **7.3.1.5. Hibakategóriák és utószerkesztési ajánlások**

Az érthetőséget vizsgáló kérdőív segítségével sikerült a gépi fordításra jellemző hibakategóriákat felállítani. Ezek a kategóriák jó kiindulópontot jelentettek a további hibaelemzéshez és az utószerkesztési elvek megállapításához. Az utószerkesztett korpusz elemzésével beigazolódt, hogy valóban ezek a kezdetben felállított hibák gyakran előfordulnak, és javítási arányuk is elég nagy. Az is kiderült azonban, hogy ezek a kategóriák nem elegendőek az elemzéshez. A korpusz kézi elemzése után a kategóriákat sikerült pontosítani, egyes kategóriákat

kibővíteni, egyeseket fel kellett osztani, valamint a egyéb kategóriákat is hozzáadni az elemzéshez.

A hibák elemzésekor csak néhány helyen utalunk az eredeti forrásnyelvi szövegre vagy annak jellegzetességeire, hiszen a cél az, hogy az eredeti szöveg nélkül tudjunk rendszerességeket találni az utószerkesztési műveletekben és a hibákban, ennek segítségével pedig automatikusan vagy kézzel tudjuk elvégezni az utószerkesztést.

Az újonnan kialakított kategóriák a következők:

1. Referenciális hiba, kontextusba beleillik, nyelviileg helyes vagy ránézésre kitalálható  
Azok a hibák kerültek ebbe a kategóriába, amelyek teljesen félrevezetik az olvasót. Ránézésre a mondat vagy szerkezet jónak, vagy nagy biztonsággal érthetőnek tűnik, de ez félrevezető.
  - (15a) A kutatók országokat kértek 40,000 31-es embertől (GF\_061)
  - (15b) A kutatók 31 országban kérdezték 40,000 embertől (USZ\_061)
2. Mondat szintű referenciális és nyelvi hiba  
A teljes mondatra ható hibák, amelyben a mondatrészek teljesen összekeverednek, a mondat referenciális jelentése (ha van) megváltozik, és általában nem kitalálható. A hiba legtöbbször tagmondat szinten is átnyúlik. Ezeket a hibákat a legnehezebb javítani, de feltétlenül szükséges. Sokszor okozza ezt a hibát az, ha egy FNY-i szónak több, más szófajú jelentése van, és másik szófaj, ráadásul másik jelentéssel kerül bele a szövegbe. A legtöbb gond az igék okozzák, mert felborítják az állítmányi viszonyokat, ezért erre külön kategóriát hoztunk létre: a 27-es kategória.
  - (16a) Az üzenet népek van igazságos, hogy gondos dolog, hogy. (GF\_030)
  - (16b) Az üzenet az embereknek az, hogy óvatosak legyenek. (USZ\_030)
3. Nyelvi és lexikai hiba: mondatrészek szerepe nem tiszta + szórend helytelen, zavaros  
Ez a kategória hasonló 2-es kategóriához, mert ez is a teljes mondatot érinti. De a hiba itt sokszor megmarad tagmondat szinten, valamint a mondat referenciális jelentése alapvetően megmaradna, ha a mondat nyelviileg helyes lenne, a szófajok megfelelőek lennének, és a

mondatrészek szerepe tiszta lenne. Ezt a hibát könnyebb javítani, és feltétlenül szükséges is.

- (17a) Még akkor is a hasznok csak rákokba, amiket okoztak alkohol és az, hogy olyan azokként dohányoztak a bélben, torokban és szájban, jelentkezhetnek. (GF\_064)
- (17b) Még akkor is a hasznok csak olyan rákok esetén érvényesülnek, amiket az alkohol és dohányzás okoztak mint például a bélben, torokban és szájban. (USZ\_064)

4. Lexikai szintű referenciális hiba + szó/kifejezés (hang)alakja teljesen eltér az odaillő kifejezésétől és jelentésmezőjük egyáltalán nem fed át Ezeket a hibák egyszerűen azok az esetek, amikor egy többjelentésű szónak nem a megfelelő jelentése kerül a mondatba. A FNY-i szöveg segítségével ezek a hibák viszonylag könnyen javíthatóak, és feltétlenül javítandóak is.

- (18a) Az önkénteseket akkor megkérték, hogy gondoljanak a **csipeszekre** megint és a pszichés számológép **alakult** jól az, hogy melyikre gondoltak. (GF\_062)
- (18b) Az önkénteseket akkor megkérték, hogy gondoljanak a **klipekre** megint, és a 'médium' számológép **kitalálta** azt, hogy melyikre gondoltak. (USZ\_062)

5. A kontextus miatt nyelvi hibából referenciális hiba lesz: a referenciális jelentés kikövetkeztethető lenne, de a kontextus és a szórend vagy az egyéb hibák miatt (pl. helytelen vonzatok, hiányzó ragok) miatt félreérthető vagy nem értelmezhető

Ez a hiba túlságosan speciális, kérdéses, hogy érdemes-e használni. Automatikus utószerkesztésre egyáltalán nem, de kézi utószerkesztéshez esetleg hasznos útmutatás lehet.

- (19a) Az ő képei egy **hajó** a gyerekekkel **felszínre került**, ahogy ez kitudódott a pár kapcsolata **ütötte a sziklákat**, miután egy párról együtt csináltak egy filmet ki szakítás. (GF\_052)
- (19b) Képei, ahol a gyerekekkel van egy hajón, felszínre kerültek, amikor kitudódott, hogy a pár kapcsolata szikláknak ütközött, miután egy párról együtt csináltak egy filmet akik szakítanak. (USZ\_052)

6. Szófajtévesztés: zárt lexikai halmazt érintő – nyelvi hiba referenciális hibát okoz

Ez a hiba általában egyszerűbb nyelvi hiba, és nagyon gyakran kifejezetten a vonatkozó névmásokat vagy kötőszavakat érinti: helyettük kérdő névmás szerepel a szövegben, vagy . Ha a mondatban más hiba nincs, akkor nem feltétlenül okoz megértési problémákat, és az utószerkesztők sem mindig javítják. Ha viszont a mondat például

kérdő szórendű, vagy más hibákat is tartalmaz, a hiba mondat szintű referenciális hibát okoz. Mivel a hibát egyszerű javítani, a feldolgozási erőfeszítés csökkentése érdekében célszerű mindig javítani.

a) nem javított

- (20a) A kutatásnak most speciális gyümölcsökön és zöldségeken kellene többet összpontosítania élesen – tartalmazó lycopene\* paradicsomokban **melyik**, a tanulmányok javasolnak, segítségük véd prosztatárák ellen – és csökkenő dohányzáson és túlsúlyosságon. (GF\_64)
- (20b) A kutatásnak most speciális gyümölcsökre és zöldségekre kellene többet élesen összpontosítania – ide értve a lycopene-t a paradicsomokban, **melyik**, a tanulmányok javasolnak, segítség a prosztatárák elleni védelemben – és a dohányzás csökkentésére és a túlsúlyosságra. (USZ\_064)

b) javított

- (21a) Ez a kor annál, hogy Moira Stewart **melyik** BBC bemondót veszttette el ellentmondásosan, a munkája volt igények közepette, hogy őt félretették mert túl öreg volt. (GF\_061)
- (21b) Ez a kor **amelyiknél** Moira Stewart BBC bemondót veszttette el ellentmondásosan a munkáját, olyan állítások közepette, hogy őt félretették mert túl öreg volt. (USZ\_061)

7. Nyelvi hiba: helytelen toldalékolás (főképp rag) miatti referenciális hiba

Vannak olyan nyelvi hibák, amelyek nem okoznak referenciális hibát, mert a CNY-i szerkezet nem létező szerkezet, és a megfelelő vonzat, szórend, stb. könnyen kikövetkeztethető. (14-es kategória) De ez az eset olyan, amelyben a ránézésre hasonló nyelvi hiba miatt más vagy nem értelmezhető lesz a referenciális jelentés. Ezeket a hibákat mindenképpen javítani kell.

- (22a) Miután miközben visszatért a szabadságról **Guernseynek** (GF\_058)
- (22b) Miután visszatértek szabadságukról **Guernsey-be** (USZ\_058)

8. Nyelvi hiba: igeidő, mód tévesztése, elhagyása vagy pl. ige helyett igenév, melléknév helyett igenév + segédige használata

Az ilyen hibák általában nem súlyosak, de vannak esetek, amikor a hibás megoldás is érthető lenne, így pedig referenciális hibát eredményeznek. Ezeket az eseteket az 1-es kategóriába átsorolhatjuk. Igeidő felcserélésekor gyakori az az eset, amikor függő beszédben az elbeszéltek ideje nem jó, például múlt idő a jelen idő helyett. Ha hiba



nem súlyos, a mondat jól értelmezhető, akkor a minimális utószerkesztéshez nem kell javítani.

a) nem súlyos hiba

- (23a) Az eredmények megmutatták, hogy **enni** további 200 g gyümölcsöt és zöldségeket egy nap alatt csökkentette a rák teljes kockázatát 3 százalékkal. (GF\_064)
- (23b) Az eredmények megmutatták, hogy **enni** további 200 g gyümölcsöt és zöldségeket egy nap alatt csökkentette a rák teljes kockázatát 3 százalékkal. (USZ\_064)

b) súlyos hiba – 1-es kategóriába sorolható!

- mód
  - (24a) Az az üzenet, **nem megy egyedül**. (GF\_030)
  - (24b) Az az üzenet, hogy **ne menj egyedül**. (USZ\_030)
- igeidő
  - (25a) Mióta hallottam, hogy **én terhes voltam**, én voltam, hiányozzon Mum több mint valaha. (GF\_051)
  - (25b) Mióta hallottam, hogy **én terhes vagyok**, anyukám jobban hiányzik, mint valaha. (USZ\_051)

9. Referenciális hiba, de a szó alakja, jelentése felidézi a megfelelő szót, és a kontextusból felidézhető a kép vagy jelenet

Ez a hiba a 4-es kategóriához nagyon hasonló, de itt a választott és odaillő szavak/kifejezések jelentésmezőjében átfedés van, még ha nem is nevezhetők szinonimáknak. A 4-es kategóriában a FNY-i szövegben nagy valószínűséggel homoníma szerepel, itt pedig valószínűleg csak poliszémiairól van szó. Hasonló a 12-es kategória is, ott pedig valamilyen mértékben már inkább szinonimákról van szó, amelyek inkább stilisztikai hibát eredményeznek. Ebben, a 9-es kategóriában szereplő hibákat az utószerkesztő megítélheti, hogy javítja-e vagy nem, mivel a lexikai hiba javítása általában könnyű. Így az utószerkesztő eldöntheti, hogy a feldolgozási erőfeszítést is mennyire növeli a javítás.

- (26a) Egy forrás Mendeshez közel mondás által magyarázta meg a **felbomlást**: (GF\_052)
- (26b) Egy forrás Mendeshez közel megmagyarázta a **szakítást**, azt mondta: (USZ\_052)

10. Lexikai hiba: nem létező szó vagy szóösszetétel használata

A fordításban sokszor előfordul, hogy olyan szóösszetételek keletkeznek, amelyek a magyarban nem léteznek. Ezeknek egy része jól érthető, másik része pedig javításra szorul. Az érthető szóösszetételeket nem kell feltétlenül javítani, bár kétségtelen, hogy hozzájárulnak a szövegek sajátságos stílusához. De mivel javításuk sokszor

nagymértékű beavatkozást igényelne, azaz akár a teljes mondatot át kellene alakítani, és az utószerkesztőtől esetleg nagyobb munkát igényel, mint amekkora mértékben a feldolgozási erőfeszítést csökkentené, inkább a változatlanul hagyást javasoljuk. Más a helyzet viszont akkor, ha a szóösszetétel nem érthető. Ilyenkor feltétlenül javítani kell.

a) Nem javított

- (27a) A 37 éves ember és 33 éves nő testeit tegnap este megtalálták, miután a rendőrök reagáltak egy zavarra az első **emelet-tulajdonnál** Cheshuntben, Herts. (GF\_032)
- (27b) A 37 éves ember és 33 éves nő testeit tegnap este megtalálták, miután a rendőrök reagáltak egy zavarra az első **emelet-tulajdonnál** Cheshuntben, Herts megyében. (USZ\_032)

b) Javított

- (28a) egy **levegőlégikisasszony** kérdezte őket, hogy üljön a repülőgép különálló részeiben (GF\_058)
- (28b) egy **légikísérő** megkérte őket, hogy üljenek a repülőgép különböző részeibe (USZ\_058)

c) Feltétlenül javítandó

- (29a) jelentés követi a Daily Mailt sikeres Banish a Bag kampány, ami vágáshoz vezetett a pazarló **hajadon-használat-szállító-zsákoknak** a számában, amik ki voltak osztva kiskereskedők. (GF\_063)
- (29b) A jelentés követi a Daily Mail sikeres Banish a Bag kampányát, ami ahhoz vezetett, hogy kevesebb pazarló **egyszer használatos szatyrokat** osztottak ki a kereskedők. (USZ\_063)

11. Stilisztikai hiba, tükörfordítás

A szövegekben nagyon gyakori a tükörfordítás. Ezeket az utószerkesztők egy része javította, de sokan nem javították. Úgy vettük észre, hogy a stilisztikai hibák javításában az a tendencia, hogy ha a mondatban egyéb hiba is van, akkor hajlamosabbak a stilisztikai hibát is javítani ezzel együtt.

a) Javított

- (30a) Mióta súlyt vesztettem el, amit a doktorom felfedezett, **nekem van egy szívhiányosságom**. Úgy tűnik, hogy nekem éveken keresztül **lehetett ezem** de azt elrejtette a méretem. (GF\_058)
- (30b) 'Mióta lefogytam, a doktorom felfedezte, hogy **szívproblémám van**. Úgy tűnik, hogy **már évek óta van**, csak a súlyom miatt nem lehetett észrevenni.' (USZ\_058)

b) Nem javított

- (31a) **Bizonyíték nem volt megtalált**, hogy ezek a töredékek kárt okoznak, de sem bizonyíték, amit tesznek, nem, a szerzők összeadtak. (GF\_063)
- (31b) **Bizonyíték nem volt megtalált**, hogy ezek a töredékek kárt okoznak, de sem bizonyíték, hogy nem, a szerzők hozzáadták. (USZ\_063)

12. Lexikai hiba, nem a megfelelő jelentés választása, de a szó jelentése hasonló az odaillőhöz, ezért ezt a hibát inkább már stilisztikai hibának tekinthetjük. A hiba javítása, ugyanúgy, mint a 10-es kategóriában, az utószerkesztő megítélésén múlhat.

Javított és nem javított:

- (32a) Arra ösztönözte Merseyside Police-t, hogy reagáljon a gonosz vándorárus ellenőrzésével kapcsolatos kérdésekre, miután azt felfedték vártak miközben kilenc hónappal korábban hoztak egy **nemzeti riadót**, amikor vándorárust, aki feltételes szabadlábra helyezésen volt még mindig, nem találhatnának meg egy **forgalmi sértéssel** kapcsolatban. (GF\_030)
- (32b) Arra ösztönözte Merseyside Police-t, hogy reagáljon a gonosz Champan ellenőrzésével kapcsolatos kérdésekre, miután kiderült, hogy kilenc hónapot vártak mire **nemzeti körözést** adtak ki, amikor Chapmant, aki feltételes szabadlábra helyezésen volt még mindig, nem találták meg egy **forgalmi sértéssel** kapcsolatban. (USZ\_030)

13. Nyelvi hiba: más szófaj, de a referenciális jelentés nem változik

Bár az eredetileg elemzett szövegben ezt a hibát nem minősítettük súlyosnak, ebből következően pedig javítandónak, a nagyobb mennyiségű szöveg elemzése után kiderült, hogy a legtöbb esetben ezeket a hibákat is javítani kell, még akkor is, ha az alapvető referenciális jelentés ugyanaz.

a) Javított

- (33a) Néhány zuhanyozik a messzi északban és keletben de a legtöbb hely **szárad**. (GF\_013)
- (33b) kevés záporosó a messzi északon és keleten de a legtöbb helyen **száraz**. (USZ\_013)

14. Nyelvi hiba: hiányzó vagy helytelen rag, névutó, vonzat, stb., ahol a helyes megoldás könnyen kikövetkeztethető

Sok ilyen hiba szerepel a szövegekben, de a mondatban levő többi hiba miatt a következtetés nehézkes lehet. Az utószerkesztők valószínűleg ezért javították ki ezeknek a hibáknak a nagy részét, és ez a stratégia javasolható is, mivel a hiba javítása egyszerű, a feldolgozási erőfeszítés

pedig lecsökkenti. Ezt a kategóriát tehát célszerű összevonni a 7-es kategóriával.

a) Javított

- (34a) Azt nem vélik annak bárki más az incidens idején a három ágyas otthonban volt és egy kézi lőfegyvert visszaszereztek a **színhelyből**. (GF\_032)
- (34b) Azt nem vélik, hogy bárki más az incidens idején a három ágyas otthonban volt és egy kézi lőfegyvert lefoglaltak a **színhelyen**. (USZ\_032)

b) Nem javított

- (35a) A döntéshez jönni nehéz, de pozitív volt miközben kényszerített engem, hogy arra gondoljak, hogy mit akarok az **életemből**, és arra vetett pillantásra, hogy nekem mi van már. (GF\_038)
- (35b) A döntést meghozni nehéz volt, de pozitív, miközben kényszerített engem, hogy arra gondoljak, hogy mit akarok az **életemből**, és arra nézzek, hogy nekem mim van már. (USZ\_038)

Az új kategóriák:

15. Az eredeti angol szó maradt a szövegben

A szótárban nem szereplő szavak a fordításban megmaradnak, csillaggal jelölve. Ezeket mindenképp javítani kell, ahogy utószerkesztők minden esetben meg is tették. Sajnos ezeknek a javítása elég nagy teher az utószerkesztőknek, mert általában utána kell nézni a szavaknak.

- (36a) A Sorrell szafaripillantásán levő tetők többsége szereti több darabka elhanyagolt füves vidéket vagy a **buddleia\*** és **willowherb\*** zsebeit, amik a második világháború villámháborúi után brit városokat jelöltek meg. (GF\_066)
- (36b) A Sorrell szafarijában levő tetők többsége úgy néz ki, mint több darabka elhanyagolt füves mező, vagy **nyáriorgona** és **fűzike** ágyások, amik a második világháború villámháborúi után a brit városokat jellemezték. (USZ\_066)

16. Angol tulajdonnév maradt a szövegben (pl. intézménynév, filmcím, földrajzi név), amelyet le kellett volna fordítani

Az utószerkesztők ezekben az esetekben nagyon sokszor nem fordították le az angol tulajdonneveket. Ez a hiba nem egészen egyértelmű, mert az olvasó esetleg ismerheti az eredeti filmcímet is, ráadásul egyes intézménynevek angol megfelelője elterjedt, míg másokat kötelezően le kellene fordítani. Ráadásul mivel ezek reáliák, a fordításuk amúgy is sok problémát vet fel. Talán a megoldás az lehet,

hogy az intézménynevekről, földrajzi nevekről listát készítünk, amelyben az utószervező könnyen megtalálja a megoldást. A film- és egyéb címeket pedig mindig le kell fordítani, esetleg zárójelben megadni az eredetit.

Ennek a kategóriának egy változata az, ha angol köznévi marad a szövegben tulajdonnévként, csillag jelölés nélkül. Ezeket feltétlenül javítani kell.

- (37a) A **Brokeback Mountain** csillag két nagybácsija azt akarja az apjától, hogy ejtsék le a birtoka végrehajtójaként mert, Kim Ledgernek van annak egy történelme, hogy pénzt pazaroljanak, mondják. (GF\_007)
- (37b) A "**Túl a barátságon**" sztárjának két nagybátyja azt szeretné, ha az apját felmentenék az örökség kezelése alól, mert mint mondják, Kim Ledger a pénzszerzésben már történelmet írt. (USZ\_007)

17. Referenciális hiba: a tulajdonnév fordítása került a szövegbe, ahol meg kellett volna maradnia az eredetinek

Ez a hiba az előzőnek pontosan a fordítottja, és általában személynéveknél fordul elő, vagy ha a szó tipikusan nem tulajdonnévi pozícióban van. Mindig javítani kell.

- (38a) A **vándorárus** akkor felfedte az igaz azonosságát és megerőszakolt és meggyilkolt Ms Hallt mielőtt kirakta a testét egy árokban. (GF\_044)
- (38b) **Chapman** akkor felfedte az igaz azonosságát és megerőszakolta és meggyilkolta Ms Hallt mielőtt kirakta a testét egy árokba. (USZ\_044)

18. Referenciális hiba: idióma, kollokáció, szókapcsolat szó szerinti fordítása

Ezek a szó szerinti fordítások attól függően jelentenek referenciális hibát, hogy a CNY-en az adott szókapcsolat létezik-e vagy nem, vagy legalább a FNY-i és a CNY-i olvasóban ugyanazt a kontextuális hatást váltja-e ki.

- (39a) Britanniát a következő néhány nap folyamán ütni fogják viharos szelek és heves eső, ahogy a tél egy **fullánkkal elindul a farkában**. (GF\_057)
- (39b) Britanniát a következő néhány nap folyamán viharos szelek és heves eső fogják érni, ahogy a tél elmúlik **egy fullánkkal a végén**. (USZ\_057)

19. Nyelvi hiba miatti referenciális hiba: a mondat alanya nem megfelelő

A mondatok alanya úgy is megváltozhat az utószerkesztésben, hogy nem feltétlenül hibajavítás történt, csak az aktív cselekvő változik meg. Az ilyen eseteket stilisztikai hibának vesszük, és javításuk opcionális. A fordított szövegekben viszont sokszor valóban a mondat alanya nem a valódi alany. Az ilyen hibákat mindenképpen javítani kell. Sokszor az ilyen mondatokban más hibák is vannak, és nagyon gyakran több alany is szerepel, a valódi alany kiléte pedig nehezen deríthető ki. Ez a fajta hiba akár besorolható a 2-es számú kategóriába, mert a mondatrészek szerepe a mondatokban nem tiszta, és a hiba a teljes mondatra hat. Sajnos az ilyen hibák javítása nehéz, sok műveletet igényel, de kötelező.

a) stilisztikai javítás

- (40a) **Ez** elszállásolhat két repülőgép legénységet egy időnél, és el van látva hordozható munkaállomásokkal az tud vezérlőállás és robottevékenységek. (GF\_012)
- (40b) Egy időben **a legénység két tagja** szállásolható el benne, és el van látva hordozható munkaállomásokkal, amikkel vezérelni lehet az állomás és a robotok tevékenységeit. (USZ\_012)

b) referenciális hiba

- (41a) A fő ok én vagyok hordani akarja a csuklyámat fent és megkértem egy olyan vallást, ami megenged engem, hogy tegye azt. (GF\_059)
- (41b) A fő ok, hogy hordani akarom a csuklyámat és van egy olyan vallásom, ami megengedi nekem, hogy tegyem azt. (USZ\_059)

20. Stilisztikai hiba: hiányzó vagy felesleges határozott vagy határozatlan névelő

A névelők hiánya nem jelent súlyos hibát a szövegekben, de mivel javításuk nagyon egyszerű, az utószerkesztők általában kijavították. A szövegekben a határozott névelők hiánya a gyakoribb, míg határozatlan névelőből van több felesleges.

- (42a) A színész bűnös csínyt tagad miközben fenyegeti, és megtámadja Brooke Mueller Sheent, 32, Aspenben amíg ők szabadságon voltak. (GF\_049)
- (42b) A színész tagadja a bűnös tettet, a fenyegetést és támadást Brooke Mueller Sheen (32 éves) ellen, Aspenben, amíg szabadságon voltak. (USZ\_049)

21. Stilisztikai hiba – az aktuális tagolásra vonatkozó szokatlan sorrend

A mondatok aktuális tagolását az utószerkesztők vegyesen javították. Ha a szövegben nincs, vagy csak kevés, nem súlyos egyéb hiba van, akkor az adott és új információ azonosítható akkor is, ha nem a megfelelő pozícióban vannak. A javítással kapcsolatban hasonlóak a

tapasztalataink, mint a 11-es kategóriánál: ha a mondatban mást is javítani kellett, akkor a mondatrészek sorrendjét is inkább megváltoztatták, mint ott, ahol mondathoz nem (nagyon) kellett hozzányúlni.

- (42a) **Az elülső rész nagyon lassan mozogni fog**, úgyhogy el tudunk várni egy párt csinos esős napoktól az országban keresztül. (GF\_057)
- (42b) **A front nagyon lassan fog mozogni**, úgyhogy elvárhatunk egy pár csinos esős napot az egész országban. (USZ\_057)

22. Stilisztikai hiba – igeikötő elválasztása vagy összekapcsolása

Az igeikötő elválasztása vagy összekapcsolása is a mondat aktuális tagolásának megváltoztatása miatt történik legtöbbször. Ez a javítás is opcionális.

- (43a) Birmingham azt mondta, hogy **rendőrséggel működnek együtt**, de hangsúlyozták, hogy a tudakozódások egy harmadik félre vagy a pártokra vonatkoznak összefüggéstelen a klubba. (GF\_006)
- (43b) Birmingham azt mondta, hogy **együttműködnek a rendőrséggel** de hangsúlyozták, hogy a vizsgálat egy harmadik félre vagy felekre vonatkozik, nincs összefüggésben a klubbal. (USZ\_006)

23. Stilisztikai hiba: mellékmondat helyett egyszerű mondatrész (pl. jelző) használata

Ez a kategória a tükörfordítás egy alkategóriája, javítására a minimális utószerkesztéshez nincs szükség.

- (44a) **A lesifotósok, akik üldözik a Mercedest**, okozták a katasztrófát? (GF\_004)
- (44b) **A Mercedest üldöző lesifotósok** okozták a katasztrófát? (USZ\_004)

24. Nyelvi hiba: hiányzó vagy felesleges utalószó és vonatkozó névmás vagy kötőszó

Ez a hiba eredetileg a 6-os kategóriába tartozó hiba, de mivel gyakran fordul elő, hasznosnak ítéltük különválasztani. A főmondatban levő, helyesen használt utalószó és a mellékmondat megkönnyíti a feldolgozási erőfeszítést, így hasznos kijavítani az ilyen hibákat, egyes esetekben pedig kötelező.

- (45a) Ez van mert nekik nekik melegnek és napfénynek kell kitéve lenniük hogy boruljanak ki. (GF\_063)
- (45b) Ez azért van, mert melegnek és napfénynek kell kitéve lenniük **ahhoz**, hogy le bomoljanak. (USZ\_063)

## 25. Számokat érintő hibák

A számok nagyon sok gondot okoznak az utószerkesztéskor, mivel gyakran az egész mondatot érintő hibákat eredményeznek. A számokat nagyon nehéz az egyes mondatrészekhez kötni, és a mondatban egészen meglepő helyeken bukkannak fel. Az utószerkesztő csak a FNY-i szöveg segítségével tudja kijavítani a hibát, és a javításnak pontosnak kell lennie. Ezeket a hibákat mindenképpen ki kell javítani.

- (46a) Az átlagos kor házasságnál nőknek emelkedett **36.5**-re emberekért és **33.8**-ért. (GF\_018)
- (46b) Az átlagos házasulandó kor férfiaknál **36.5**-re a nőknél **33.8**-ra emelkedett. (USZ\_018)

## 26. Helytelen vagy hiányzó egyeztetés

Előfordul, hogy az egyeztetés (bármilyen típusú: személy, szám, egyéb ragok) esetleg teljesen elmarad, vagy pedig halmazott mondatrészeknél először megvan, de utána eltűnik. Ez a fajta hiba elég ritka, az érthetőségre általában nincsenek nagy hatással. Javításuk viszont egyszerű, tehát ajánlott.

- (47a) Mi annyira sokkal társaságkedvelőbbek és **aktívak** vagyunk, ez kedvel nekünk van az élet egy új haszonbérlete. (GF\_058)
- (48b) 'Sokkal társaságkedvelőbbek és **aktívabbak** lettünk, mintha tényleg új életünk lenne. (USZ\_058)

## 27. Nyelvi hiba miatti referenciális hiba: igéket érintő szófajcsere

Az ilyen jellegű hibák kijavítása nehéz feladat, sokszor sok szerkesztési művelettel, akár az egész mondat átszerkesztésével vagy átírásával járhat. Kellő gyakorlattal a rendszerességek felfedezhetőek, de általában a javításhoz a FNY-i szöveget is fel kell használni.

- (49a) Azt mondta, hogy az élet testileg nehéz volt azelőtt egyszerű feladatokkal **szeress** felsétálni a lépcsőkre és arra, hogy befűzzenek cipőfűzőket, amik bizonyítanak egy kihívást. (GF\_058)
- (49b) Azt mondta, hogy az élet "fizikailag nehéz" volt azelőtt, olyan egyszerű feladatok is, **mint** felsétálni a lépcsőn vagy befűzni a cipőfűzőt is kihívást jelentettek. (USZ\_058)



28. Referenciális hiba helyetti stilisztikai hiba: nem megfelelő igeikötő használata

Meglepően sokszor fordul elő, hogy az eredeti és utószerkesztett szövegekben ugyanaz az ige más igeikötővel szerepel, vagy az igeikötő nélküli ige igeikötővel, vagy igeikötős ige igeikötő nélkül. Az igeikötő jelentése a szövegkörnyezet nélkül így akár egészen más is lehet, de mivel a szövegekből általában könnyen kikövetkeztethető a megfelelő jelentés, az utószerkesztésnek nincs olyan nagy jelentősége.

- (50a) Egy olyan idős özvegy, aki nem akarta azt, hogy **elnevezzék**, azt mondta, hogy a folyóparton sétáltatta a kutyáját, amikor látta a katasztrófát és száguldott hogy segítséget kapjon. (GF\_011)
- (50b) Egy idős özvegy, aki nem akarta azt, hogy **megnevezzék**, azt mondta, hogy a folyóparton sétáltatta a kutyáját amikor látta a katasztrófát és száguldott hogy segítséget kapjon. (USZ\_011)

29. Hiányzó mondatrész

Ha a FNY-i szövegben olyan ellipszis van, amely a CNY-ben nem elfogadható, ott a hiányzó mondatrészt legalább névmással pótolni kell.

- (51a) **Ha ez nem**, a szabályokat, amik megakadályozzák, hogy női repülőgép legénységek szolgáljanak az amerikai tenger alatti flottában, automatikusan emelni fogják. (GF\_039)
- (51b) **Ha ez nem történik meg**, a szabályokat, amik megakadályozzák, hogy női legénységek szolgáljanak az amerikai tenger alatti flottában, automatikusan megszüntetik. (USZ\_039)

30. Nyelvi hiba: alanyi – tárgyas ragozás összekeverése

Az alanyi és tárgyas (határozott és határozatlan, tárgyas és általános) ragozás összekeverésére van példa, a leggyakoribb a *mondott* – *(azt) mondta, összeadott – hozzátette* páros. A szövegtípus miatt ezek annyira gyakoriak, hogy az olvasó néhány mondat után „megtanulja” és így megérti őket. Ha viszont a szövegeket „tapasztalatlan” olvasónak szerkesztjük, akkor érdemes ezt a hibát is kijavítani.

- (52a) Alan és Jan Coupe - ki nyerte a becsületet miután 15-nél többet hullatott, kő közöttük - **mondott** elveszteni a súlyt megmentette az életüket. (GF\_058)
- (52b) Alan és Jan Coupe - aki kivívta a megtiszteltetést, miután több, mint 15 stone-t fogytak - **mondta**, hogy a fogyás megmentette az életüket. (USZ\_058)
- (53a) Orava, ki **operált meg** Barcelona főnököt, Pep Guardiolát, és sportolókat szintén Haile Gebrselassie és Merlene Ottey, **mondott**: (GF\_055)

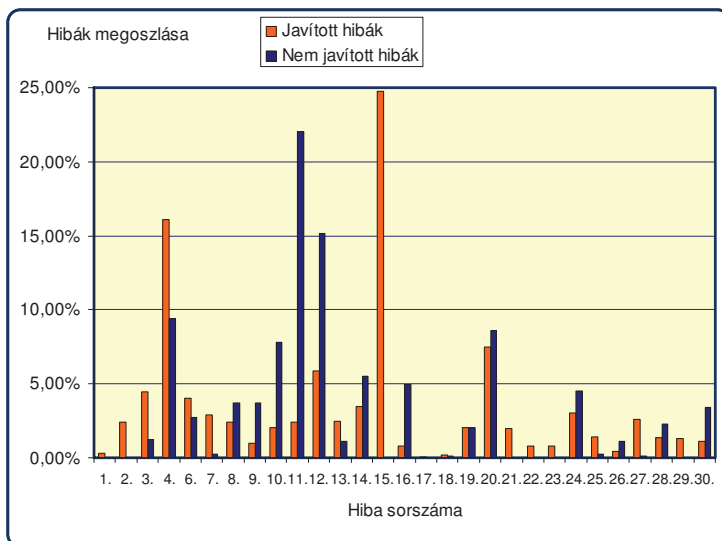
- (53b) Orava, aki **megoperálta** a barcelonai főnököt, Pep Guardiolát is, és sportolókat Haile Gebrselassiet és Merlene Otteyt, **azt mondta:** (USZ\_055)

A listát látva úgy tűnhet, hogy a kezdeti kategórialista egyáltalán nem volt elégséges, mivel az új kategóriák száma jóval nagyobb, mint az eredeti kategóriáké, majdnem kétszer annyi. De ha megnézzük a hibák súlyosságát és a szövegekben előforduló hibák arányát, akkor láthatjuk, hogy a kezdeti kategóriák sokkal súlyosabbak és jóval gyakrabban fordulnak elő, mint a későbbiekben észrevett hibák, ráadásul az új kategóriák sokszor a már meglevő kategóriák alesetei.

A hibák előfordulási arányát a korpuszban a következő táblázat tartalmazza. Az első oszlop a hiba sorszámát jelenti a fenti listában, a második oszlop azt jelzi, hogy a hibák aránya a szegmensek számára vonatkozik vagy a tokenekre. A teljes mondatokat vagy tagmondatokat érintő hibák arányát a szegmensekből számítottuk ki, a szavakat, szóösszetételeket érintőket pedig a tokenek számából. A harmadik oszlopban az adott hibakategóriában levő javított hibák száma szerepel, a negyedikben pedig ezeknek az aránya az összes javított hibához képest. Az ötödik oszlop a javított hibák aránya abban a kategóriában az összes hibához képest, a hatodik a nem javított hibák száma, a hetedik a nem javított hibák aránya az összes nem javított hibához képest, a nyolcadik az adott kategóriában a nem javított hibák aránya az összes hibához képest. A kilencedik oszlop az összes hiba abban a kategóriában, a tizedik pedig ennek az aránya az összes szegmenshez vagy tokenhez képest, attól függően, hogy a hiba milyen szintet érint. Néhány kategóriában csak a javított hibákra vonatkozóan vannak adatok, ezek az adatok ki vannak húzva, és azok a cellák üresek, ahol így az arányosításnak nincs értelme. A táblázatban levő értékekben elképzelhető egy kis mértékű tévedés, mert a kézi elemzéshez sok az adat, és az elemzés hosszú és fáradságos.

Hiba	Mon- dat vagy szó	Javi- tott darab	Javi- tottak aránya	Javított/ Nem javított arány	Nem javi- tott darab	Nem javított aránya	Nem javított /javított t arány	Össze- sen	Össz százalék
1.	sz	18	0,33%	100,00%	0	0,00%	0,00%	18	0,07%
2.	m	134	2,43%	100,00%	0	0,00%	0,00%	134	10,18%
3.	m	246	4,46%	95,72%	11	1,24%	4,28%	257	19,53%
4.	sz	887	16,07%	91,44%	83	9,38%	8,56%	970	3,69%
6.	m	223	4,04%	90,28%	24	2,71%	9,72%	247	18,77%
7.	sz	159	2,88%	98,76%	2	0,23%	1,24%	161	0,61%
8.	m	134	2,43%	80,24%	33	3,73%	19,76%	167	12,69%
9.	sz	56	1,01%	62,92%	33	3,73%	37,08%	89	0,34%
10.	sz	112	2,03%	61,88%	69	7,80%	38,12%	181	0,69%
11.	m	134	2,43%	40,73%	195	22,03%	59,27%	329	25,00%
12.	sz	323	5,85%	70,68%	134	15,14%	29,32%	457	1,74%
13.	sz	137	2,48%	93,20%	10	1,13%	6,80%	147	0,56%
14.	sz	191	3,46%	79,58%	49	5,54%	20,42%	240	0,91%
15.	sz	1365	24,74%	100,00%	0	0,00%	0,00%	1365	5,20%
16.	sz	44	0,80%	50,00%	44	4,97%	50,00%	88	0,34%
17.	sz	3	0,05%	100,00%	0	0,00%	0,00%	3	0,01%
18.	m	9	0,16%	90,00%	1	0,11%	10,00%	10	0,76%
19.	m	113	2,05%	86,26%	18	2,03%	13,74%	131	9,95%
20.	sz	412	7,47%	84,43%	76	8,59%	15,57%	488	1,86%
21.	m	110	1,99%	-	-	-	-	-	-
22.	sz	45	0,82%	-	-	-	-	-	-
23.	m	43	0,78%	-	-	-	-	-	-
24.	m	167	3,03%	80,68%	40	4,52%	19,32%	207	15,73%
25.	m	78	1,41%	97,50%	2	0,23%	2,50%	80	6,08%
26.	m	25	0,45%	71,43%	10	1,13%	28,57%	35	2,66%
27.	sz	143	2,59%	99,31%	1	0,11%	0,69%	144	0,55%
28.	sz	76	1,38%	79,17%	20	2,26%	20,83%	96	0,37%
29.	sz	70	1,27%	-	-	-	-	-	-
30.	sz	61	1,11%	67,03%	30	3,39%	32,97%	91	0,35%
Összesen:		5518		86,18%	885		13,82%	6403	
A hiányzók nélkül:		5250		85,57%	885		14,43%	6135	

31. táblázat: A javított és nem javított hibák



**6. ábra :A javított és nem javított hibák aránya**

A diagramon a javított és nem javított hibák arányát látjuk az összes javított illetve nem javított hibához képest. Ezeket az adatokat láthatjuk a táblázat 4. illetve 7. oszlopában. Az adatokból azt láthatjuk, hogy melyik a leggyakoribb illetve legritkább javított hiba, valamint ugyanezt a nem javított hibáknál. A két kategória összehasonlításával vigyázni kell, mert a diagram nem az egymáshoz képesti arányokat mutatja. A leggyakoribb javított hiba a 15-ös (eredeti angol), és ez a legritkább nem javított hiba. A második leggyakoribb javított hiba a 4-es, a feltétlenül javítandó lexikai hiba, és érdekes, hogy ez a hiba nem javított hibák között is elég nagy arányú. A nem javított hibák között a leggyakoribb a 11-es és 12-es, a tükrörfordítás és a (nem súlyos) lexikai hiba. Ezek aránya a valóságban még ennél is nagyobb lehet, mert a megítélésük szubjektív, és könnyű elsiklani felette.

A táblázatból és a diagramon láthatjuk, hogy a legnagyobb számú kijavított hiba az eredeti, 1-14-es kategóriában van. Kivétel ez alól az 1-es kategória, amelyben viszonylag kicsi a hibák száma, valamint a 15-ös kategória, amelyben a legtöbb javított hiba van, és amely egyértelműen javítandó.

Eredetileg az abszolút minimális utószerkesztéshez az 1-8-as kategóriákba tartozó hibákat kijavítását határoztuk meg. Ezt alátámasztja az, hogy a korábban a nem javítandó kategóriába eső 10-es, 11-es és 12-es kategóriában van a legtöbb nem javított hiba. Ezenkívül a 4-es kategóriában is sok a nem javított hiba, de ezt már korábban feltétlenül kijavítandó hibaként határoztunk meg, és mivel a kijavított hibák száma is itt nagy, ezt nem is változtatjuk meg. A 20-as kategóriában (névelők) nagy a javított és a nem javított hibák aránya is. A névelőhibák nem súlyos hibák, de a javításokból úgy látszik, hogy az elfogadhatóságra mégis hatással vannak. Ezért és mivel könnyen javíthatóak, ezt a kategóriát a javításra ajánlott (de nem kötelező) kategóriába soroljuk. A 24-es kategóriában (utalószavak, névmások) is elég nagy a hibák száma mindkét kategóriában, ezért ezt is a javításra ajánlott hibák közé vesszük fel. Az utószerkesztési ajánlásokat kategóriák szerint a 7.3.2 részben mutatjuk be.

#### **7.3.1.6. A hibák hatása az elfogadhatóságra**

A szövegek elemzése és a hibakategóriák megállapítása után úgy gondoljuk, hogy a stilisztikai hibákat az érthetőség szempontjából nem kell hibának tekinteni, ezek inkább a feldolgozási erőfeszítést növelik, így csökkentik az elfogadhatóságot. Az elfogadhatóság szubjektivitását bizonyítja, hogy szinte minden kijavított hibatípusra van ellenpélda, tehát ugyanolyan jellegű hiba, amely nincs kijavítva, ami az olvasók különböző tolerancia-szintjét, hozzáállását tükrözi, és azt jelzi, hogy a különböző utószerkesztők különböző szinten javítanak. Nézzünk erre egy példát:

Nem javított melléknévi igenév + segédige:

(54a) A hosszú életük miatt ez lehetséges, hogy a Galapagosön levő legöregebb teknősbékák közül néhány **élő volt**, amikor Charles Darwin 1835-ben látogatóban volt. (GF\_008)

(54b) A hosszú életük miatt ez lehetséges, hogy a Galapagosön levő legöregebb teknősbékák közül néhány **élő volt**, amikor Charles Darwin 1835-ben látogatóban volt. (USZ\_008)

Javított melléknévi igenév + segédige:

(55a) Jan és én mindig együtt cselekedtünk dolgokat és **gondterhelt volt** az ha nem tettem valamit a súlyomról, én nem lehetek ott hosszú ideig, tehát beleegyeztem, hogy járjak vele. (GF\_058)

(55b) 'Jan és én mindig együtt csináltunk dolgokat és **aggódott**, hogy ha nem csinálok valamit a súlyommal, én nem lehetek ott hosszú ideig, tehát beleegyeztem, hogy eljárjak vele.' (USZ\_058)

Az, hogy az utószerkesztők stílusa egy kicsit különböző, segít megtalálni a középútat az utószerkesztésben. Vannak, akiket jobban zavarnak a stilisztikai hibák, és vannak, akik ezeket egyáltalán nem javítják ki, sőt, nyelvi hibákat is a szövegben hagynak, ha ez az érthetőséget számukra nem befolyásolja. De míg ez a sokszínűség előny a jelenlegi felmérésben, a valódi utószerkesztési feladatokban hátrány lehet. Ez pontosan alátámasztja Allen (Allen 2001: 26) állítását, hogy szükség van utószerkesztési előírára, iránymutatásra, amely az utószerkesztők és a megrendelőik dolgát is megkönnyíti. Az utószerkesztőknek segít a döntéshozatalban, a megrendelő pedig tudja, hogy mit várhat, így viszonylag egységes szövegeket kaphat, valamint az utószerkesztésre fordított idő is jobban kiszámítható.

## 7.3.2. Utószerkesztési minták

### 7.3.2.1. Bevezetés

Az utószerkesztési műveletek és a szövegek hosszú és kitaró tanulmányozása után megállapíthatjuk, hogy a téma, terület megadása, leszűkítése nemcsak a gépi fordításban, hanem az utószerkesztésben is nagy előnyt jelenthet. A 4-es számú lexikai hibák javítása úgy részben automatizálható lenne, ha megadjuk, hogy témák alapján a lexikai elemet hogyan kell javítani. Ugyanez az utószerkesztési

útmutatóban megkönnyítené az utószerkesztők dolgát. A másik lehetőség a részletes szabálymegadás kontextussal, ekkor nem a témát adjuk meg, hanem azokat az egyedi eseteket, amikor ezeket az elemeket cserélni kell. Nagy valószínűséggel mindkét módszerre szükség van, mert vannak esetek, ahol a téma nehezen adható meg, vagy a csere inkább a közvetlen kontextustól függ, és nem a témától. Mindkét módszer hátránya az, hogy munkaigényes, valamint hosszú listákra és szabályhalmazokra van szükség hozzájuk. Az automatikus feldolgozásban a hosszú listák nem annyira hátrányosak, de kétséges, hogy egy utószerkesztő számára ezek valódi segítséget jelentenek-e.

Nézzünk meg két példát: *emberek – férfiak*: Az *ember* szó (vagy bármilyen toldalékolt alakja) 81-szer szerepelt az eredeti szövegekben, ebből 27-szer kellett *férfira* cserélni, minden olyan esetben, ahol vagy a szó közvetlen környezetében, vagy 1-2 mondaton belül szerepelt a *nő* szó. Ha ezt a tudást felhasználjuk, akkor alkothatunk szabályokat az – automatikus vagy kézi – utószerkesztéshez.

A másik, nagyon jellemző példa: *elülső – front*, vagy akár említhetjük a *zuhaný – zápor* párt is.

Az időjárással kapcsolatos cikkek alapján megfogalmazhatjuk azt a szabályt, hogy a ha a szöveg témája az időjárás, akkor az *elülső* szót *front*-ra kell cserélni. Természetesen előfordulhat, hogy még az időjárási témájú cikkekben sem kellene a cserét elvégezni, de ennek valószínűsége sokkal kisebb, tehát egy ilyen szabállyal több hibát ki tudunk javítani, mint amennyit okozunk. Egy kellően nagy minta alapján megállapítható egy ilyen szabály jósága, azaz annak a valószínűsége, hogy milyen arányban javít, és ha ez a valószínűség egy kívánt mérték alatt van, akkor a szabály elfogadható lesz. Mivel a mintánk ehhez túlságosan kicsi, nem tudjuk ezeket a valószínűségeket hitelesen megállapítani, de ez nem is a célunk. A cél most az, hogy rendszerességeket találjunk a javításokban, és hogy az elveket, valamint a javításokhoz kötődő szabályokat egyáltalán meg tudjuk fogalmazni.

### 7.3.2.2. Szabályok a hibakategóriák alapján

A fenti elemzések tükrében szeretnénk utószerkesztési szabályokat és mintákat adni, amelyek főleg a kézi utószerkesztéshez adhatnak segítséget.

A kötelezően javítandó hibák (a fenti hibalista alapján):

1. Referenciális hiba, kontextusba beleillik, nyelvtanilag helyes vagy ránézésre kitalálható
2. Mondat szintű referenciális és nyelvi hiba
3. Nyelvi és lexikai hiba: mondatrészek szerepe nem tiszta + szórend helytelen, zavaros
4. Lexikai szintű referenciális hiba + szó/kifejezés (hang)alakja teljesen eltér az odaillő kifejezésétől és jelentésmezőjük egyáltalán nem fed át
5. Szófajtévesztés: zárt lexikai halmazt érintő – nyelvi hiba referenciális hibát okoz
6. Nyelvi hiba: helytelen toldalékolás (főképp rag) miatti referenciális hiba
7. Nyelvi hiba: igeidő, mód tévesztése, elhagyása vagy pl. ige helyett igenév, melléknév helyett igenév + segédige használata, de csak ha súlyos hiba (1-es kategóriába sorolható)
9. Nyelvi hiba: más szófaj, de a referenciális jelentés nem változik
15. Az eredeti angol szó maradt a szövegben
16. Angol tulajdonnév maradt a szövegben (pl. intézménynév, filmcím, földrajzi név), amelyet le kellett volna fordítani
17. Referenciális hiba: a tulajdonnév fordítása került a szövegbe, ahol meg kellett volna maradnia az eredetinek
19. Nyelvi hiba miatti referenciális hiba: a mondat alanya nem megfelelő, de csak ha referenciális hiba (ld. 19. b. pont)
24. Nyelvi hiba: hiányzó vagy felesleges utalószó és vonatkozó névmás vagy kötőszó
25. Számokat érintő hibák



26. Helytelen vagy hiányzó egyeztetés

27. Nyelvi hiba miatti referenciális hiba: igéket érintő szófajcsere

29. Hiányzó mondatrész

A következő hibákat nem feltétlenül kötelező, de az elfogadhatóság érdekében ajánlott kijavítani:

14. Nyelvi hiba: hiányzó vagy helytelen rag, névutó, vonzat, stb., ahol a helyes megoldás könnyen kikövetkeztethető

20. Stilisztikai hiba: hiányzó vagy felesleges határozott vagy határozatlan névelő

28. Referenciális hiba helyetti stilisztikai hiba: nem megfelelő igekötő használata

A következő hibák kijavítása az utószerkesztő ítéletére van bízva. Ezekben az esetekben csak az érthetőséget kell mérlegelni, hogy az utószerkesztő munkát ne növeljük feleslegesen. Ha az utószerkesztő érthetőnek találja az érintett szövegrészt, akkor nem kell ezeket a hibákat kijavítani.

9. Referenciális hiba, de a szó alakja, jelentése felidézi a megfelelő szót, és a kontextusból felidézhető a kép vagy jelenet

10. Lexikai hiba: nem létező szó vagy szóösszetétel használata

11. Stilisztikai hiba, tükörfordítás

12. Lexikai hiba, nem a megfelelő jelentés választása, de a szó jelentése hasonló az odaillőhöz, ezért ezt a hibát inkább már stilisztikai hibának tekinthetjük. A hiba javítása, ugyanúgy, mint a 10-es kategóriában, az utószerkesztő megítélésén múlhat.

18. Referenciális hiba: idióma, kollokáció, szókapcsolat szó szerinti fordítása

19. Nyelvi hiba miatti referenciális hiba: a mondat alanya nem megfelelő, stilisztikai javítás (ld. 19. a.)

21. Stilisztikai hiba – az aktuális tagolásra vonatkozó szokatlan sorrend

22. Stilisztikai hiba – igekötő elválasztása vagy összekapcsolása

A következő hibákat egyáltalán nem kell javítani, kivéve, ha esetleg egyéb hibák miatt ezek javítása egyszerűbb és érthetővé teszi a mondatot.

23. Stilisztikai hiba: mellékmondat helyett egyszerű mondatrész (pl. jelző) használata

30. Nyelvi hiba: alanyi – tárgyas ragozás összekeverése

### 7.3.2.3. Egyéb utószerkesztési minták és szabályok

A következő listában a fent említett hibakategóriákon kívüli javítási szabályokat adunk meg, valamint olyan mintákat, szabályokat, amelyek megkönnyíthetik az utószerkesztők munkáját annyiban, hogy az ismétlődő, az érthetőséget megnehezítő elemeket könnyebben felismerik. Nem minden szabály kötelező érvényű, vannak, amelyek inkább csak segítséget adnak.

#### 1. Szabályok

- A birtokos névmásokat be kell építeni a ragokba.
- A személyes névmásokat a legtöbb esetben be kell építeni a ragokba.
- A párbeszédeket, film- és egyéb címeket idézőjelek közé kell tenni.
- A vesszőket az alárendelő mellékmondatok határán ki kell javítani.
- A mértékegységeket át kell váltani és magyarul rövidítve megadni
- Ha idézetet vezet be vagy zár le, akkor a *hozzátesz, mond, ír* igéket tárgyas ragozásúra kell változtatni, esetleg utalószóval bevezetni (azt mondta) Ilyenkor általában az ige nagyon rövid tagmondatban van, tárgya pedig egy másik mondat vagy tagmondat. A tagmondatot gyakran kettőspont zárja.
- A *csinál* igét a tárgynak megfelelően át kell írni. (pl. *panaszt csinál, igényt csinál, érkezést csinál* stb.)

#### 1. Minták

- Szavak, lexikai elemek cseréje

A következő táblázat példákat tartalmaz azokra a lexikai elemekre, amelyeket a korpusz alapján másik lexikai elemre kell cserélni. Az első oszlopban van a törlendő elem, a másodikban a helyette beszúrandó elem, a harmadik oszlop a lexikai elem szöveggörnyezetét, a negyedik pedig a szöveg témáját (ha meghatározható) tartalmazza.

Mit	Mire	Kontextus	Téma
ember	férfi	nők	nők és férfiak összehasonlítása

Mit	Mire	Kontextus	Téma
zuhany	zápor		időjárás
összead	hozzátesz tárgyas ragozás	múlt idő, önmagában egy tagmondat, kettőspont utána, a tárgy másik tagmondatban van	idézet jelzésekor
igény	állítás	nincs előtte/utána <i>iránt</i> névutó	
csipogások	üzenetek	Twitter	
csattanás	fénykép		fényképezés
dátum	randi, randevú		párkapcsolatok
keltez	randevúzik, találkozik		párkapcsolatok
egyetlen	egyedülálló	férfi, nő	párkapcsolatok
nincs út	sehogy, semennyire, egáltalán		
elkövet	elvesz	tárgya egy személy	párkapcsolatok
elszalaszt	hiányol, eltűnik, hiányzik belőle		
elülső	front		időjárás
emel	felold, megszüntet	tárgya törvény, szabály	
ölelkezik	támogat	tárgya elvont, személytelen	
igazságos	csak	összekevert mondatrészek	
szeret, szeress, kedvel	mint, mintha	tagmondatok között, az alany tipikusan nehezen meghatározható	hasonlításakor, példák említésekor
megcímez	foglalkozik valamivel	tárgya elvont fogalom	
lovász	vőlegény		párkapcsolatok
otthon	lakás	?	?
színhely	helyszín		minden, kivéve színházi, filmes téma
alakok	számok, adatok	azt mutatják,	

Mit	Mire	Kontextus	Téma
esedezett nem bűnös	ártatlannak/ nem bűnösnek vallotta magát		
termelés	forgatás	film, show, darab stb.	filmek, színház
hamarosabban	hamarabb/korábban		
mondja sajnálkozó	bocsánatot kér		
ivadék	olajban süített (ételek)	ételek	ételek, evés
közöttük	együtt, együttesen		
súlyt veszít	fogy		
vesztes súly	fogyás		
kikopik	viseli		
témák	alanyok	cselekvő állítmány	kísérletek, felmérések
kiborul	lebomlik	tárgya valamilyen anyag, test stb.	
a mások	a többiek		
utolsó éjszaka	tegnap este		
megenged	lehetővé tesz/nyújt valamit		

32. táblázat: Ajánlott lexikai cserék

b) Egyéb, nem kötelező minták, ajánlások

- Bizonyos esetekben az *épp, éppen* jelentése *egyszerűen* vagy *csak*.
- Személyek kora után ki kell tenni, hogy *éves*.
- Az eredeti mondatban minden elemnek lesz megfelelője az utószekszett mondatban, de ezek sokszor egészen más helyen és más formában vannak. Ezt az elvet szem előtt tartva könnyebb a mondatot megérteni, majd átrendezni és átírni.
- Középfokú melléknév helyett sokszor több + alapfokú melléknév kell. (pl. *nőtlenebb férfiak* → *több nőtlen férfi*)
- *Ha* kötőszóval kezdődő mondatokban feltételes mód lehet szükséges a kijelentő mód helyett.

A fent látható szabályok és példák listája nem teljeskörű, hiszen még a kis méretű, megvizsgált korpuszban is található további szabályosságokat még részletesebb vizsgálatokkal. A fenti lista kezdőpontja lehet a további, nagyobb korpuszon végzett felmérésnek.

#### **7.4. Az eredmények összefoglalása**

Egy kérdőív segítségével létrehoztunk a géppel fordított szövegekhez kezdeti hibakategóriákat, amelyeket súlyoztunk az érthetőségre gyakorolt hatásuk alapján. Így létrejött egy javítási index, amelynek értéke [-5;+5] között van. Az index alapján meghatároztuk, hogy a negatív javítási indexszel rendelkező hibákat az abszolút minimális utószerkesztéshez feltétlenül ki kell javítani. További szövegek elemzése alapján kiderült, hogy a minimális utószerkesztéshez további hibák kijavítására van szükség. Ehhez a korábbi hibalistát kiegészítettük, majd meghatároztuk, hogy ezek közül melyek szükségesek vagy ajánlottak a minimális utószerkesztéshez, és melyek azok, amelyeket az utószerkesztő ítéletére kell bízunk.

A szövegek elemzése után további szabályokat és utószerkesztési mintákat határoztunk meg. A minták és szabályok létrehozásában a cél az volt, hogy a szövegek érthetővé, és így elfogadhatóbbá is váljanak. Reményeink szerint idővel az érthetőség elégséges kritérium lesz a gépi fordítás elfogadhatóságához.

## 8. Az utószerkesztés hatása az érthetőségre és az elfogadhatóságra

### 8.1. Bevezetés

Allen et al. (Allen 2001: 27) azt hiányolja, hogy nincs egységes követelményrendszer és szintmeghatározás a minimális utószerkesztésre vonatkozóan, és pontosan ennek a hiánynak a pótlását célozza meg. De a nyelvek szerkezeti különbségeiből eredően ez a szempontrendszer nem lehet egyforma minden nyelvre. A nyelvek különbségei miatt eleve a gépi fordítás hibái is más jellegűek, ráadásul ezek másképp befolyásolják a megértést is. Tehát a magyarra fordított szöveg minimális utószerkesztési követelményei nem lehetnek ugyanazok, mint az angolra fordított szöveg követelményei.

A hibák jellegét egyébként nemcsak a nyelv, hanem a fordítórendszer működési elve is meghatározza; más jellegű hibákat produkál például egy szabályalapú és egy statisztikai alapú rendszer, valamint a hibák függenek a rendszer „háttértudásától” is, azaz a mintáktól, a szabályoktól, a nyelvtantól, szótártól, korpuszoktól, területspecifikus rendszernél a területtől stb. Tehát azt is mondhatjuk, hogy az egy rendszert alapul vevő utószerkesztési modell nagy valószínűséggel csak ahhoz a rendszer lesz megfelelő.

A 7. fejezetben bemutatott kérdőíves vizsgálat eredményeképpen sikerült meghatározni egy minimális utószerkesztési követelményrendszert az adott szövegekhez. Az elemzésben minden elem kapott egy indexet az érthetősége alapján. Az index 5 és -5 között van, az 5-ös indexű teljesen érthető, a -5-ös pedig teljesen félreértelmezhető. Ezek alapján kialakult egy rendezett lista, amely a hibákat tartalmazza súlyosságuk alapján. A listában 0 és negatív indexű elemekben szereplő hibát tekintjük a leg súlyosabb hibának, azaz olyanoknak, amelyeket feltétlenül ki kell javítani.

## 8.2. A minimális utószerkesztés vizsgálata kérdőívvel

### 8.2.1. A szövegek érthetősége

A fentiek alapján létrehoztunk egy abszolút minimális utószerkesztési követelményrendszert, és kijavítottuk az előzetesen már megvizsgált szövegeket. A minimálisan utószerkesztett szövegekből ugyanazokkal a kérdésekkel elkészítettük a kérdőívet. A rövid kérdéseket tartalmazó rész ebbe a kérdőívbe nem került bele. A kérdőívet 49 fő töltötte ki, szintén nagyrészt BSc szintű egyetemi hallgatók, de nem ugyanazok, mint akik az előző, nyers gépi fordításról készült kérdőívet.

Az eredmények a következőképpen alakultak:

	<b>Teljes</b>	<b>Hiányos</b>	<b>Jó és helyte- len</b>	<b>Nem tudja</b>	<b>Nagy- részt helyte- len</b>	<b>Hely- telen</b>	<b>Index</b>
1.	79,6%	14,3%	4,1%	0,0%	0,0%	2,0%	0,893
2.	87,8%	12,2%	0,0%	0,0%	0,0%	0,0%	0,969
3.	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	1,000
4.	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	1,000
5.	83,7%	2,0%	2,0%	2,0%	2,0%	8,2%	0,760
6.	8,2%	73,5%	4,1%	0,0%	4,1%	10,2%	0,510
7.	89,8%	6,1%	0,0%	0,0%	2,0%	2,0%	0,908
8.	91,8%	0,0%	2,0%	0,0%	2,0%	4,1%	0,867
9.	77,6%	18,4%	0,0%	4,1%	0,0%	0,0%	0,913
	79,8%	14,1%	1,4%	0,7%	1,1%	2,9%	0,869

33. táblázat: Utószerkesztett szövegek érthetőségének relatív értékei

A kérdések teljes egészében megegyeztek az előző kérdőív kérdéseivel. Az indexek szöveges jelentéseit itt láthatjuk:

	<b>Kérdés</b>	<b>Index</b>	<b>Szövegesen kifejezve</b>
1.	<b>Miért került bajba David Beckham?</b>	0,893	érthető
2.	<b>Mi a baj a drótkerítéssel?</b>	0,969	érthető
3.	<b>Mi történt Freddie Williamsszel és a lányával?</b>	1,000	érthető
4.	<b>Mit vettek el tőlük?</b>	1,000	érthető
5.	<b>Mennyit?</b>	0,760	nagyrészt érthető +
6.	<b>Mi lett velük?</b>	0,510	nagyrészt érthető -
7.	<b>Miről szól az RAC felmérése?</b>	0,908	érthető
8.	<b>Mit tesznek a nők, ha eltévednek?</b>	0,867	nagyrészt érthető +
9.	<b>Mit tesznek/milyenek a férfiak, ha eltévednek?</b>	0,913	érthető
	<b>Összes szöveg:</b>	<b>0,869</b>	nagyrészt érthető +

**34. táblázat: Az abszolút minimálisan utószerkesztett szövegrészek érthetőségi indexe**

Az abszolút minimálisan utószerkesztett szövegek (melléklet 11.3) eredményeiből azt láthatjuk, hogy ha a leg súlyosabb hibákat kijavítjuk a nyers gépi fordításban, akkor a szövegek érthetősége ugrásszerűen megnő, átlagban alig tér el az emberi fordítás eredményeitől, amelynek átlagos indexe 0,974, ami 0,105, azaz nagyjából 10% különbséget jelent, ha itt is csak 0-tól számítjuk jónak a szövegeket. A [-1;+1] skálán ez viszont már csak 5%. Az utószerkesztett szöveg így a nagyrészt érthető kategóriájának felső részébe, az emberi fordítás pedig az érthető kategóriába esik. Az eredeti, nyers fordítás index értéke 0,414 volt, amely a félig érthető + kategóriába esett, ehhez képest a minimális utószerkesztéssel óriási javulást érünk el. Ezeket a szövegeket valóban csak az abszolút minimális utószerkesztési elvek alapján javítottuk, tehát a negatív javítási indexszel rendelkező hibákat.

Ezekből a kísérleti eredményekből látható, hogy érdemes a minimális utószerkesztéssel foglalkozni, hiszen az emberi fordítás érthetőségétől csak 5%-



ban tér el, időben mérve pedig a minimális utószerkesztés csak töredékét teszi ki az emberi fordításnak.

### **8.2.2. Szubjektív vélemények az utószerkesztett szövegekről**

Miután az alanyok válaszoltak a kérdésekre, valamint látták az emberi fordításokat is, megkérdeztük azt is, hogy mi a véleményük az utószerkesztett szövegekről. A válaszban saját szavaikkal írhatták le a véleményüket a szövegek érthetőségéről, olvashatóságáról, szerkezetéről, stílusáról. Az eredmények a következőképpen alakultak:

4% nem válaszolt a kérdésekre egyáltalán.

10-10% azt válaszolta, hogy az utószerkesztett szövegek nehezen illetve jóval nehezebben érthetőek. 10% szerint nagyjából érthetőek, 15% pedig azt írta, hogy a lényegét ki lehet venni. 28% írta, hogy nehezebben érthetőek, megértésükhöz több idő kell, 17% pedig többszöri olvasásra megértette. 25,53% szerint a minimálisan utószerkesztett szöveg alig valamivel kevésbé érthető, 15% pedig azt mondta, hogy érthető. 6% emelte ki, hogy elveszik információ, 2% (1 fő) szerint sok a félrefordítás. 11% emelte ki a nyelvtani furcsaságokat, 15% pedig a stílusbeli különbségeket. 4% hangsúlyozta a szövegek közti különbséget is. 15% írta, hogy sok esetben a szavak nem megfelelő jelentése kerül a szövegbe. Érdekes, hogy 2% (1 fő) írta, hogy az utószerkesztett szöveg azért nehezebben érthető, és azért kell hozzá több idő, mert kevésbé megszokott. A számok itt természetesen nem tesznek ki 100%-ot, mert a kategóriák nem egymást kizáróak. Az egymást kizáró érthetőségi kategóriákban kapott válaszok aránya a következő táblázatban látható:

Kategória	Fő	Százalék
Érthető	1	2,13%
Szinte teljesen érthető	6	12,77%
Nagyjából/lényeg érthető	14	29,79%
Többszöri olvasásra/nagyobb erőfeszítéssel érthető	10	21,28%
Nehezen érhető	11	23,40%
Nem értelmezhető válasz	3	6,38%
Nem válaszolt	2	4,26%

**35. táblázat: A vélemények megoszlása az érthetőségről**

Ezek alapján 66% érthetőnek tartja a minimálisan utószerkesztett szövegeket, 23,4% pedig nehezen érthetőnek. Az olvashatósággal (gördülékenységgel) kapcsolatos vélemények a következők:

Kategória	Fő	Százalék
Jól olvasható	1	2,13%
Kicsit nehezebben olvasható	3	6,38%
Nehezebben olvasható, többszöri olvasás, nagyobb figyelem szükséges	8	17,02%
Nehezen olvasható	16	34,04%
Szokatlan stílus	5	10,64%
Nem írt az olvashatóságról	12	25,53%
Nem válaszolt	2	4,26%

**Táblázat: A vélemények megoszlása az olvashatóságról**

Az összes válaszadó közül csak 8%-nak volt jó véleménye az olvashatóságról, láthatjuk, hogy 34% nehezen olvashatónak ítélte a szövegeket, de ehhez hozzáadhatjuk azoknak a számát is, akik szerint nagyobb erőfeszítés kell az olvasáshoz, és így 51% nehezen olvashatónak tartja ezeket a szövegeket. Ez a szám pedig már 73%, ha csak azokat vesszük figyelembe, akik az olvashatóságról adtak véleményt, ami egy nagyon nagy arány.

Az olvashatóság rossz eredményei mellett nem csodálkozhatunk azon, hogy milyen sokan (21%) emelték ki azt, hogy a megértéshez, olvasáshoz nagyobb erőfeszítés, többszöri olvasás és/vagy hosszabb idő kell, amikor az érthetőséggel kapcsolatos véleményüket írták le (akár nehezen, akár jól érthetőnek vélték a

szöveget). A rossz olvashatóság ellenére viszont a vélemények általánosságban pozitív hozzáállást és nagy engedékenységet tükröznek.

Az eredmények egyértelműen azt mutatják, hogy a rossz olvashatósághoz képest a szövegek jelentősen érthetőbbek a vártnál. A legérdekesebb eredmény az, hogy annak ellenére, hogy az érthetőségi index mennyire magas volt a szövegeknél, tehát a fontos információt megértették, az olvasóknak mégis csak összesen 66%-a érezte érthetőnek. Ez is azt bizonyítja, hogy az érthetőség és az elfogadhatóság nem azonos. Az olvasó szereti megérteni a szöveg minden részletét, és zavarja, ha egy kis részlet is hiányzik, valamint az is, ha erőfeszítéseket kell tenni.

### 8.3. A minimális utószerkesztés vizsgálata automatikus módszerrel

Láttuk tehát, hogy az abszolút minimális utószerkesztéssel a szövegek érthetősége ugrásszerűen javul. Vajon az elfogadhatóságuk mennyiben javul? Ha az abszolút minimálisan utószerkesztett szöveget minősítjük az automatikus, elfogadhatóságot mérő programmal, akkor a következő eredményeket kapjuk:

Szöveg	Elfogadhatóság	Érthetőségi index	Érthetőségi index [0;1]
Nyers gépi fordítás	0,358	0,414	0,707
Abszolút minimálisan utószerkesztett szövegek	0,391	0,869	0,934
Minimálisan utószerkesztett szövegek	0,524	-	-
Teljesen utószerkesztett szövegek/ emberi fordítás	0,872	0,974	0,987

36. táblázat: Különböző mértékben utószerkesztett szövegek elfogadhatósága

A harmadik oszlopban az érthetőségi indexeket látjuk. Kérdés, hogy a két indexet össze lehet-e hasonlítani, hiszen az értékelés intervalluma nem azonos. Ezért a negyedik oszlopban az érthetőségi index [0;1] intervallumra leképezett értékét láthatjuk, amelyet már fenntartások nélkül össze lehet hasonlítani az elfogadhatósági indexszel. A harmadik sorban szereplő minimálisan

utószerkesztett szövegek az új, kiegészített, szigorúbb alapelvek alapján utószerkesztett szövegeket jelentik, de ezekre nem készült az érthetőségre vonatkozó kérdőíves felmérés. A teljesen utószerkesztett szövegek és az emberi fordítások nagyon hasonlítanak egymásra. Ezek már teljes mértékben emberi szövegnek tekinthetőek, de az utószerkesztett változatok még közelebb állnak a gépi fordításhoz, jobban megtartják a sorrendeket, szókincset és szerkezeteket.

#### **8.4. Az eredmények összefoglalása**

A minimálisan utószerkesztett szövegekről az olvasóknak jobb a véleménye, mint az első és a jelenlegi vizsgálat alapján a géppel fordított szövegekről. Az első vizsgálat alapján 72% adott negatív véleményt, a minimálisan utószerkesztett szövegekkel kapcsolatban pedig 34%-nak volt negatív véleménye. Az érthetőségi vizsgálat itt is azt bizonyította, hogy a szövegek érthetőbbek, mint ahogy azt az olvasók megítélik, és az elfogadhatóság alapján elvárthoz képest az érthetőségi index nagyon magas. A automatikus mérések alapján is bebizonyosodott, hogy bár az elfogadhatóság növekszik a javítással, de ez a növekedés nem áll egyenes arányban az érthetőség növekedésével. Az érthetőség még az abszolút minimális utószerkesztésnél is majdnem megközelíti a teljes megértést, és az indexe több, mint 2 tizeddel javul. Az elfogadhatóság eközben csak 3 század értékkel javul, és még mindig rendkívül alacsony.

## 9. Összegzés

### 9.1. A tézisek összegzése

#### 1. tézis

A minőség két fontos szempontja, az elfogadhatóság és az érthetőség a géppel fordított szövegek esetében kapcsolatban áll ugyan, de egymással nem megegyező, és nem felcserélhető fogalom. Azt állítjuk, hogy a géppel fordított szövegek elfogadhatóságának mértéke jóval kisebb, mint ahogy azt az érthetőségük indokolná.

Egy kérdőíves vizsgálattal felmértük a géppel fordított szövegek érthetőségét. Az érthetőség mérésére létrehoztunk egy  $[-1;+1]$  közötti érthetőségi indexet, amely a géppel fordított szövegek érthetőségére 0,414, amely a félig érthető kagetóriát jelenti. A vizsgálat alapján úgy ítéltük, hogy az olvasók a szövegek alapvető információ-tartalmát ki tudják szűrni, és az első látásra teljesen érthetetlennek tűnő szöveget is meglepően jól megértik. Ennek ellenére a szövegekről a véleményük elvárásaikhoz képest rossz: 72%-uknak negatív a véleménye. A szövegek elfogadhatósága nagyon alacsony.

#### 2. tézis

Azt állítjuk, hogy az elfogadhatóság szubjektív kritérium, és számos tényezőtől függ. Ennek ellenére az elfogadhatóságot is mérhetjük objektív módon úgy, hogy ehhez kellően sok olvasói mintát használunk. Az objektív méréshez egy automatikus mérési módszert hoztunk létre.

Arra a megállapításra alapozva, hogy az emberi szövegekhez való hasonlóság növeli a géppel szövegek elfogadhatóságát (Amigó et al. 2006: 22-23), létrehoztunk és megvalósítottunk egy automatikus minősítő módszert, amellyel a szövegek elfogadhatóságát értékeljük a  $[0;1]$  intervallumban. A minősítéshez használt referencia-szövegeket az interneten található magyar szövegek adják. A módszer jól korrelál az emberi ítélettel, valamint valóban különbséget tesz a géppel

fordított, utószerkesztett és az emberi szövegek között, és kimutatja a szövegek javulását. Az automatikus minősítés előnye, hogy kis különbséget és kis mértékű javulást is ki tud mutatni. A minősítő módszer újszerűsége az, hogy nincs hozzá szükség az értékelendő szöveg emberi fordítására.

3. A gépi fordítás első minőségi szempontja az elfogadhatóság, amely egyrészt minden más minőségi szempont felett áll. Ha a szöveg nem elfogadható, az egyéb minőségi szempontok sem lesznek relevánsak. Az általunk vizsgált géppel fordított szövegek érthetősége a vártnál magasabb, de ahhoz, hogy a szövegek érthetősége érvényesüljön, és az olvasók profitálhassanak belőle, az elfogadhatóságot is növelni kell. Másrészt az elfogadhatóság olyan szempont, amely az egyéb minőségi szempontoktól függ. Az elfogadhatóságot kétféle módszerrel növelhetjük:

### **3.1. tézis**

A szöveg utószerkesztésével: a hibák kijavításával, azaz a szövegek minimális utószerkesztésével, amellyel az érthetőségük az elfogadhatóságnál még nagyobb mértékben javul.

A géppel fordított szövegek minimális utószerkesztésével érthetőségük a 0,414-es értékről 0,869-re javult. A kérdőíves felmérésben ehhez képest csak 66% ítélte teljesen vagy nagyjából érthetőnek őket. Az automatikus méréssel a minimálisan utószerkesztett szövegek átlagos elfogadhatósága pedig 0,524. A géppel fordított szövegek átlagos elfogadhatósága 0,358 volt, tehát az utószerkesztés után az érthetőségben bekövetkezett javulás lényegesen nagyobb, mint az elfogadhatóság javulása.

### **3.2. tézis**

Az olvasók hozzáállásának javításával. Azt állítjuk, hogy a relevanciaelmélet a gépi fordítás folyamatát is magyarázza. A relevanciát pedig úgy is növelhetjük, hogy az olvasó hozzáállását megváltoztatjuk, méghozzá úgy, hogy hajlandó legyen nagyobb feldolgozási erőfeszítést tenni a számára szükséges információ

kinyeréséhez. Ebben a folyamatban pedig a fordítástudománynak is nagy szerepe van.

Bemutattuk, hogy a relevanciaelmélet a gépi fordítás folyamatára is magyarázatot ad. A fordítási folyamatban a fordító „jellegzetességei” okozzák feldolgozási erőfeszítés növekedését: a kognitív környezetének hiányosságai, valamint a hallgató, a beszélő és a fordító kognitív környezetének különbsége. A feldolgozási erőfeszítés növekedése csökkenti a szövegek relevanciáját, még akkor is, ha a megfelelő kontextuális hatást megkapják, aminek az indoka részben az, hogy a szokatlanság és a hibák miatt az elfogadhatóság csökken, és így kontextuális hatást kisebbnek érzik. A relevancia tehát növelhető az elfogadhatóság növelésével. A géppel fordított szövegek helyének, szerepének, használhatóságának, használati szabályainak meghatározásában a fordítástudomány sokat segíthet. Ha az olvasók elvárásait reális mederbe tereljük, a gépi fordítás elfogadhatósága, ezzel együtt pedig relevanciája is növekszik.

4. Az utószerkesztési elvek meghatározásához szükség van a szövegek elemzésére. Az elemzéshez kétféle módszert választottunk, és az utószerkesztésnek két szintjét határoztuk meg.

#### **4.1. tézis**

A géppel fordított szövegek kétségtelenül sok hibát tartalmaznak. Ezek a hibák csoportosíthatóak, és súlyosságuk meghatározható az alapján, hogy mennyire gátolják az érthetőséget. A géppel fordított szövegek hibáinak kategorizálásához nem megfelelőek az emberi fordításhoz létrehozott hibakategóriák. Azt állítjuk, hogy a sorba rendezés után legsúlyosabbnak ítélt hibák kijavításával a szövegek érthetősége nagymértékben javul, az emberi fordítást megközelíti. Ezt a javítást hívjuk abszolút minimális utószerkesztésnek.

Az érthetőséget mérő kérdőíves vizsgálat eredményei alapján meghatároztuk a géppel fordított hibáit, ezeket kategorizáltuk, majd létrehoztunk egy [-5;+5] közötti javítási indexet. A javítási index értéke azt tükrözi, hogy az adott hiba

mennyire befolyásolja az érthetőséget. A hibákra vonatkozó indexek meghatározása után a legsúlyosabb, tehát negatív indexszel rendelkező hibakategóriákat soroltuk az abszolút minimális utószerkesztési irányelvekhez.

#### **4.2. tézis**

A nagyobb mértékű, minimális utószerkesztéshez szükséges szigorúbb hibakategóriák megállapíthatóak, ha az olvasók javítják ki a szövegeket, majd a javítások alapján mintákat állapítunk meg. Az olvasói javítások automatikus kiszűrése majd kategorizálása alapján létrehozott javítási mintákkal minimális utószerkesztési alapelveket állapítunk meg.

Nagyobb mennyiségű géppel fordított szöveg utószerkesztési műveleteinek kiszűréséhez elkészítettünk egy programot. Az így kapott eredményekből valamint az utószerkesztett szövegek kézi elemzése után az abszolút minimális utószerkesztési kategóriákat kibővítettük, valamint utószerkesztési szabályokat és mintákat határoztunk meg a minimális utószerkesztés segítésére. A nyers szövegeket átlagos, fordítással nem foglalkozó olvasók javították ki.

### **9.2. A kutatás újdonságai, korlátai és kiterjesztési lehetőségei**

Bár a gépi fordítás minősítése a nemzetközi szakirodalomban rendkívül népszerű téma, a magyar kutatások inkább a fordítórendszerek, gépi fordítási módszerek tökéletesítésére összpontosítanak. Ugyanígy az utószerkesztés lehetőségeit sem vizsgálják. A kutatás tehát újdonságot jelent a következő területeken:

- A gépi fordítás elfogadhatóságát elválasztja az érthetőségtől.
- Az elfogadhatóság mérésére új, objektív, automatikus módszert ad.
- Új, módosított számítási módszert ad a töredezettségi büntetési értékref a METEOR automatikus értékelési módszerhez.
- Új, kis műveletigényű algoritmust ad egy intervallum legjobb lefedéséhez.
- Specifikus utószerkesztési javaslatokat tesz az angol-magyar, MetaMorpho által fordított nyers gépi fordítás utószerkesztésére.

A kutatás korlátai, egyben fejlesztési lehetőségei:



- Az elfogadhatóságot mérő program egyelőre csak magyar nyelvű szövegekre alkalmazható.
- Minden kutatást egyetlen fordítórendszerrel végeztünk. Az automatikus értékelési módszerrel bármilyen típusú szöveg elfogadhatósága mérhető, valamint az utószerkesztési műveleteket kivonatoló program is alkalmazható bármilyen rendszerrel fordított szöveg feldolgozására alkalmas, tehát egy következő lépésben a megadott kutatási módszerek tovább általánosíthatóak.
- A vizsgált korpusz mérete a nyelvtchnológiában alkalmazott korpuszok méretéhez képest nagyon kicsi. Az utószerkesztési műveleteket kivonatoló program használható további, nagyméretű párhuzamos korpusz feldolgozására, amelynek segítségével további specifikus mintákat és szabályokat lehet megállapítani.
- Az automatikus utószerkesztési lehetőségeket nem vizsgáltuk, de a legnagyobb továbblépési lehetőséget ebben látjuk. Az utószerkesztési alapelvek, és főképp a specifikus minták segítségével érdemes az automatikus utószerkesztés lehetőségeivel foglalkozni, vagy a tapasztalatokat a MetaMorpho továbbfejlesztésében felhasználni.
- A kijavított és nem kijavított hibák vizsgálatára és összehasonlítására a jelenlegi, utószerkesztési műveleteket kivonatoló program nem alkalmas. A nem javított hibák kiszűrése ezért csak kézzel lehetséges, ezért a jelenlegi kutatás nem is fókuszál erre kellő mértékben. A nem kijavított hibák összehasonlítása a kijavítottakkal teljesebb képet adna a feltétlenül szükséges utószerkesztési műveletekről.
- Véleményünk szerint a kontrollált nyelvi megközelítés is az elfogadhatóság növekedését segítheti elő. A hibák elemzése segítséget ad a kontrollált nyelvi szabályok és sablonok elkészítéséhez, valamint iránymutatás készíthető a felhasználóknak, hogy milyen jellegű szövegekkel, milyen szerkezetekkel, szókinccsel, stb. kaphatnak jobb minőségű fordítást.

A disszertáció tehát egy logikai vonalat követve ad irányvonalat a további, teljesebb körű kutatáshoz, amelynek segítségével az angol–magyar gépi fordítás minősége javítható, valamint tisztázza, milyen körülmények között és milyen célokra jól használható a gépi fordítás.

## 10. Irodalomjegyzék

- Agarwal, A., Lavie, A. 2008. Meteor, M-BLEU and M-TER: Evaluation metrics for high correlation with human rankings of machine translation output. In *Proceedings of the Third Workshop on Statistical Machine Translation. Association for Computational Linguistics*. Columbus, Ohio. 115–118.
- Aikawa, T., Schwartz, L., King, R., Corston-Oliver, M., Lozano, C. 2007. Impact of controlled language on translation quality and post-editing in a statistical machine translation environment. In: *Proceedings of MT Summit XI*. Copenhagen, Denmark. 1-7.
- Albrecht J., Hwa, R. 2008. The role of pseudo references in MT evaluation. In: *Proceedings of the Third Workshop on Statistical Machine Translation. Association for Computational Linguistics*. Columbus, Ohio. 187–190.
- Allen, J. 2001. Post-Editing: An Integrated Part of a Translation Software Program. In: *Language International*. Vol. 13. No. 2. 26-29.
- Allen, J. 2003. Post-editing. In: Somers, H. (ed.) 2003. *Computers and Translation: A Translator's Guide*. Amsterdam: John Benjamins Publishing Company.
- Allen, J., C. Hogan: 2000. Toward the Development of a Post-Editing Module for Raw Machine Translation Output: A Controlled Language Perspective. In *Proceedings of the Third International Workshop on Controlled Language Applications*. Seattle, WA. 62-71.
- Amigó, E., Giménez, J., Gonzalo, J., Márquez, L. 2006. MT evaluation: human-like vs. human acceptable. In: *Annual Meeting of the ACL Proceedings of the COLING/ACL on Main conference poster sessions*. Sydney, Australia. 17-24.
- Arnold, D.J., Balkan, L., Meijer, S., Humphreys, R.L., Sadler, L. 1994. *Machine Translation: an Introductory Guide*. London: Blackwells-NCC.
- Babych B., Hartley, A. 2004. Extending the BLEU MT evaluation method with frequency weightings. In: *ACL 2004: 42nd annual meeting of the Association for Computational Linguistics. Proceedings of the conference*. Barcelona, Spain. 621-628.
- Babych, B. 2005. *Information extraction technology in machine translation: IE methods for improving and evaluating MT quality*. PhD thesis. Leeds: University of Leeds, Centre for Translation Studies.
- Banerjee, S., Lavie, A. 2005. METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments. In: *Proceedings of Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization at the 43rd Annual Meeting of the Association of Computational Linguistics (ACL-2005)*. Ann Arbor, Michigan, USA. 65-72.
- Bar-Hillel, Y. 1951. The present state of research on mechanical translation. *American Documentation*. Vol 2. No. 4. 229-237.

- Beaugrande, R-A. de, and Dressler, W. U. 1981. *Introduction to text linguistics*. London: Longman.
- Bell\_Blog. Random posts, ramblings and exercises in writing, by a translator. 2010. *I feel the need, the need for... machine translation*. 2010. szeptember 6. <http://babelon.wordpress.com/2010/09/06/i-feel-the-need-the-need-for-machine-translation/>
- Bellos, D. 2010. I, Translator. In: *The New York Times*. 2010. március 20. [http://www.nytimes.com/2010/03/21/opinion/21bellos.html?\\_r=1](http://www.nytimes.com/2010/03/21/opinion/21bellos.html?_r=1) és [http://www.nytimes.com/2010/03/21/opinion/21bellos.html?pagewanted=2&\\_r=1](http://www.nytimes.com/2010/03/21/opinion/21bellos.html?pagewanted=2&_r=1)
- Boitet, C. 1988. Bernard Vauquois' contribution to the theory and practice of building MT systems: a historical perspective. In: *Second International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*. Carnegie Mellon University, Center for Machine Translation. Pittsburgh, Pennsylvania, USA. 1-18.
- Boitet, C., Blanchon, H., Seligman, M., Bellyneck, V. 2009. Evolution of MT with the Web. In: *Proceedings of the Conference "Machine Translation 25 Years On"*. Cranfield, England. 1-13.
- Bowker, L., Ehgoetz, M. 2005. A Pilot Study Exploring User Acceptance of Machine Translation. In: *CLiNE 05 3rd Computational Linguistics in the North-East Workshop*. Université du Québec en Outaouais, Gatineau, Québec. 1-4.
- Callison-Burch, C., Fordyce, C., Koehn, P., Monz, C., Schroeder, J. 2008. Further meta-evaluation of machine translation. In: *ACL-08: HLT. Third Workshop on Statistical Machine Translation, Proceedings. (ACL WMT-08)*. The Ohio State University, Columbus, Ohio, USA. 70-106.
- Callison-Burch, C., Osborne, M., Koehn, P. 2006. Re-evaluating the role of BLEU in machine translation research. In: *EACL-2006: 11th Conference of the European Chapter of the Association for Computational Linguistics*. Trento, Italy. 249-256.
- Catford, J. C. 1965. *A Linguistic Theory of Translation. An Essay in Applied Linguistics*. London: Oxford University Press.
- Chandioux, J. 1976. METEO, an operational system for the translation of public weather forecasts. In: *FBIS Seminar on Machine Translation. American Journal of Computational Linguistics*. 27-36.
- Church, K. W., Hovy, E. H. 1993. Good applications for crummy machine translation. *Machine Translation*. Vol. 8. No. 4. 239-258.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L. 2001. *Algoritmusok*. Budapest: Műszaki Könyvkiadó.
- Coughlin, D. 2003. Correlating automated and human assessments of machine translation quality. In: *Proceedings of MT Summit IX*. New Orleans, USA. 63-70.

- Dave, S. Parikh, J., Bhattacharyya P. 2001. Interlingua-based English–Hindi Machine Translation and Language Divergence. *Journal of Machine Translation*. Vol. 16. No. 4. 251-304.
- Denkowski, M., Lavie, A. 2010. Extending the METEOR machine translation evaluation metric to the phrase level. In: *NAACL HLT 2010: Human Language Technologies: the 2010 annual conference of the North American Chapter of the Association for Computational Linguistics. Proceedings*. Los Angeles, California. 250-253.
- Doddington, G., 2002. Automatic Evaluation of Machine Translation Quality using N-gram Cooccurrence Statistics. In: *Proceedings of 2nd Human Language Technologies Conference (HLT-02)*. San Diego, CA. 128–132.
- Doyon, J, Doran, C, Means, C.D, Parr, D. 2008. Automated Machine Translation Improvement Through Post-Editing Techniques: Analyst and Translator Experiments. In: *Proceedings of AMTA 2008*. Waikiki, Hawai'i. 1-8.
- Dróth J. 2001. *Formatív értékelés a fordítás oktatásában*. (Kiadatlan doktori értekezés) Pécs: Pécsi Tudományegyetem.
- Dugast, L., Senellart, J., Koehn, P. 2007. Statistical Post-Editing on SYSTRAN's Rule-Based Translation System. In: *Proceedings of the Second Workshop on Statistical Machine Translation*. Prague: Association for Computational Linguistics. 220–223.
- Dugast, L., Senellart, J., Koehn, P. 2008. Can we relearn an RBMT system? In: *Proceedings of the Third Workshop on Statistical Machine Translation* Columbus, Ohio, USA. 175-178.
- Estrella, P., Popescu-Belis, A., King, M. 2008. Improving quality models for MT evaluation based on evaluators' feedback. *LREC 2008: 6th Language Resources and Evaluation Conference*. Marrakech, Morocco.
- FEMTI - a Framework for the Evaluation of Machine Translation in ISLE. 2002. <http://www.issco.unige.ch:8080/cocoon/femti/st-home.html>
- Fiederer, R., O'Brien, S. 2009. Quality and machine translation: a realistic objective? *Journal of Specialised Translation*. Vol. 4. London. 52-74.
- Gamon, M., Aue, A., Smets, M. 2005. Sentence-level MT evaluation without reference translations: beyond language modeling. In: *10th EAMT conference "Practical applications of machine translation"*. Budapest. 103-111.
- Giménez, J., Màrquez, L. 2008a. A smorgasbord of features for automatic MT evaluation. In: *ACL-08: HLT. Third Workshop on Statistical Machine Translation (ACL WMT-08). Proceedings*. Columbus, Ohio, USA. 195-198.
- Gósy, M. 2005. *Pszicholingvisztika*. Budapest: Osiris.
- Grimalia, A., Chandioux, J. 1992. Made to measure solutions. In: Newton, J. (ed.) 1992. *Computers in Translation: A Practical Appraisal*. London: Routledge. 33-45.

- Gröbner, T. 2004. Egyértelműsítés és „mozaifordítás” a MetaMorpho rendszerben. In: Alexin Z., Csendes D., (szerk.) 2004. A 2. Magyar Számítógépes Nyelvészeti Konferencia előadásai. 88–91. Szeged: SZTE.
- Gutt, E. A. 2000. *Translation and relevance: Cognition and context*. Manchester: St. Jerome Publishing.
- Guzmán, R. 2007a. Automating MT post-editing using regular expressions. *Multilingual Computing*. #90, Vol. 18. No. 6. Idaho, USA. 49-52.
- Guzmán, R. 2007b. Manual MT Post-editing: “if it's not broken, don't fix it!” *Tranlation Journal*. 2007/10. <http://accurapid.com/journal/42mt.htm>
- Hamon, O., Mostefa, D., Arranz, V. 2008. Diagnosing human judgments in MT evaluation: an example based on the Spanish language. In: *MATMT 2008: Mixing Approaches to Machine Translation*. Donostia-San Sebastian, Spain. 19-26.
- Hatim, B., Mason, J. 1990. *Discourse and the Translator*. London: Longman.
- Hatim, B., Mason, I. 1997. *The translator as communicator*. London: Routledge.
- Heltai P. 2004. A fordító és a nyelvi norma I. *Magyar Nyelvőr*. 128. évf. 4. szám. 407-434.
- Heltai P. 2005. A fordító és a nyelvi norma II. *Magyar Nyelvőr*. 129. évf. 1. szám 30-58.
- Hovy, E., King, M., Popescu-Belis, A. 2002a. Computer-aided specification of quality models for machine translation evaluation. In: *LREC-2002: Third International Conference on Language Resources and Evaluation. Proceedings*. Las Palmas de Gran Canaria, Spain. 1239-1246.
- Hovy, E., King, M., Popescu-Belis, A. 2002b. Principles of Context-Based Machine Translation Evaluation. In: *Machine Translation*. Vol. 17. No. 1. 43-75.
- Hovy, E., King, M., Popescu-Belis, A. 2002c. An introduction to MT evaluation. In: *LREC-2002: Third International Conference on Language Resources and Evaluation. Workshop: Machine translation evaluation: human evaluators meet automated metrics*. Las Palmas, Canary Islands. 1-7
- Hutchins, J. 1995. “The whiskey was invisible”, or persistent myths of MT. *MT News International*. No. 11. 17-18.
- Hutchins, J. 2005a. Current commercial machine translation systems and computer-based translation tools: system types and their uses. *International Journal of Translation*. Vol.17. No. 1-2. 5-38.
- Hutchins, J. 2005b. Towards a definition of example-based machine translation. In: *MT Summit X. Proceedings of Workshop on Example-Based Machine Translation*. Phuket, Thailand. 63-70.
- Hutchins, J., Somers, H. 1992. *An Introduction to Machine Translation*. London: Academic Press Limited.

- Kevin Duh. 2008. Ranking vs. regression in machine translation evaluation. In: *Proceedings of the Third Workshop on Statistical Machine Translation. Association for Computational Linguistics*. Columbus, Ohio. 191-194.
- Kis B. 2008. A fordítástechnológia és az alkalmazott nyelvtudomány. (Kiadatlan doktori értekezés.) Pécs: Pécsi Tudományegyetem.
- Klaudy K. 2004. *Bevezetés a fordítás elméletébe*. Budapest: Scholastica.
- Klaudy, K. 2005. *Bevezetés a fordítás gyakorlatába*. Budapest: Scholastica.
- Knight, K., Chander, I. 1994. Automated Postediting of Documents. In: *Proceedings of the 12th National Conference on Artificial Intelligence*. Seattle, USA: American Association for Artificial Intelligence. 779-784.
- Koby, G. S., Krings, H. P. 2001. *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. Kent, USA: Kent State University Press.
- Koehn, P., et al. 2007. Moses: Open source toolkit for statistical machine translation. In: *Proceedings of the Second Workshop on Statistical Machine Translation Demo and Poster Sessions*. Association for Computational Linguistics. Prága. 177–180.
- Koehn, P., Monz, C. 2006. Manual and Automatic Evaluation of Machine Translation between European Languages. In: *Proceedings of the Workshop on Statistical Machine Translation*, 102–121.
- Lin C., Och, F. J. 2004. Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics. In: *ACL 2004: 42nd annual meeting of the Association for Computational Linguistics: Proceedings of the conference*. Barcelona, Spain. 605-612.
- Mártonyi É., Varga Á., 2007. Learners' translation – machine translation. Lexical cohesion in language learners' and machine translation. In: Lendvai E. (szerk.) 2007. *Translatologica Pannonica. A PTE BTK Fordításudományi Kutatóközpont elektronikus folyóirata*. I. évf. 1. szám. Pécs: PTE BTK. 80-89.
- Melamed, I. D., Green R., Turian, J. P. 2003. Precision and Recall of Machine Translation. In: *Proteus technical reports*. Vol. 3. No. 4. 1-3.
- Mitamura, T. 1999. Controlled Language for Multilingual Machine Translation. In: *Proceedings of Machine Translation Summit VII*. Singapore. 46-52.
- Navarro, G. 2001. A guided tour to approximate string matching. *ACM Computing Surveys*. Vol. 33, No. 1. 31-88.
- Newmark, P. 1993. *Paragraphs on Translation*. Philadelphia: Multilingual Matters.
- Newton, J. 1992. Introduction and overview. In: Newton, J. (ed.) 1992. *Computers in Translation: A Practical Appraisal*. London: Routledge. 1-13.
- Nida, E. A. 1964. *Towards a Science of Translating*. Leiden: Brill.
- Novák A., Tihanyi L., Prószték G. 2008. The MetaMorpho Translation System. Proceedings of the Workshop. In: *Proceedings of the Third Workshop on Statistical Machine Translation*. Columbus, Ohio, USA. 111–114.

- Papineni, K., Roukos, S., Ward, T., and Zhu, W. J. 2002. BLEU: a method for automatic evaluation of machine translation. In: *ACL-2002: 40th Annual meeting of the Association for Computational Linguistics*. 311–318.
- Pierce, J. R., Carroll, J. B. et al. 1966. *Language and Machines – Computers in Translation and Linguistics*. Automated Language Processing Advisory Committee (ALPAC) report. Washington, D. C.: National Academy of Sciences, National Research Council. 1-124.
- Popescu-Belis, A. 2003. An Experiment in Comparative Evaluation: Humans vs. Computers. *Proceedings of the Machine Translation Summit IX*. New Orleans. 307-314.
- Popovic, M. and Ney, H. 2007. Word error rates: Decomposition over POS classes and applications for error analysis. In: *Proceedings of ACL Workshop on Machine Translation*. Prague, Czech Republic. 48-55.
- Prósztéký .G., Tihanyi L., Ugray G. 2004. Moose: a robust high-performance parser and generator. In: *Proceedings of the 9th Workshop of the European Association for Machine Translation*. Foundation for La Valletta, Malta, International Studies. 138-142.
- Prósztéký, G. 1994. Industrial Applications of Unification Morphology. *Proceedings of the 4th Conference on Applied Natural Language Processing (ANLP)*. Stuttgart, Germany: University of Stuttgart. 157–159.
- Prósztéký G., Tihanyi L. 2002. MetaMorpho: A Pattern-Based Machine Translation Project. In: *24th Translating and the Computer Conference*. London, United Kingdom. 19-24.
- Prósztéký G., Kis B. 1999. *Számítógéppel emberi nyelven*. Bicske: Szak Kiadó.
- Pym, A. 1992. Translation Error Analysis and the Interface with Language Teaching. In: Dollerup, C., Loddegaard, A. (eds.) 1992. *The Teaching of Translation*. Amsterdam: John Benjamins. 279–288.
- Reboul, A., Moeschler, J. 2005. *A társalgás cselei*. Budapest: Osiris.
- Reifler, E. 1955. The Mechanical Determination of Meaning. In: Nirenburg, S., Somers, H., Wilks Y. (eds.) 2003. *Readings in machine translation*. Cambridge, MA, USA: MIT Press. 21-36.
- Reiß, K. 1976. *Texttyp und Übersetzungsmethode: Der operative Text*. Scriptor: Kronberg.
- Reiß, K. 1981. Type, kind and individuality of text. Decision making in translation. In: Venuti, L. (Hrsg.) 2004. *The Translation Studies Reader*. 161-171. London, New York: Routledge.
- Reiß, K. 1978. Anwendbarkeit der Texttypologie mit besonderer Berücksichtigung der Sachprosa. In: Gomard, K., Poulsen, S. (Hrsg.) 1978. *Stand und Möglichkeiten der Übersetzungswissenschaft*. Acta Jutlandica LII. Humanities Series 54. Aarhus. 27-35.



- Roberts, M. 2007. Optimizing Post-Editing. *Language Weaver*. Vol. 3. No. 5. [http://www.imakenews.com/lweaver/e\\_article000828087.cfm?x=b11,0,w](http://www.imakenews.com/lweaver/e_article000828087.cfm?x=b11,0,w)
- Norvig, P., Russell, S. 2000. *Mesterséges intelligencia modern megközelítésben*. Budapest: Panem.
- Ryan, R. M., & Deci, E. L. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*. No. 25. 54-67.
- Sager, J. 1983. Quality and standards – the evaluation of translations. In: Picken, C. (ed) 1989. *The translators handbook*. London: Aslib. 91-102.
- Sager, J. C. 1994. *Language Engineering and Translation: Consequences of automation*. Amsterdam: John Benjamins.
- Schubert, K. 1992. Esperanto as an intermediate language for machine translation. In: Newton, J. (ed.) 1992. *Computers in Translation: A Practical Appraisal*. London: Routledge. 68-95.
- Shieber, S. M. 1986. An Introduction to Unification-Based Approaches to Grammar. *CSLI Lecture Notes Series*. Vol. 4. Stanford, USA: Center for the Study of Language and Information.
- Simard, M., Goutte, C., Isabelle, P. 2007a. Statistical phrase-based post-editing. In: *NAACL-HLT-2007 Human Language Technology: the conference of the North American Chapter of the Association for Computational Linguistics*. Rochester, NY. 508-515.
- Simard, M., Ueffing, N., Isabelle, P., Kuhn, R. 2007b. Rule-based Translation With Statistical Phrase-based Post-editing. In: *Proceedings of the Second Workshop on Statistical Machine Translation*. Association for Computational Linguistics. Prague. 203-206.
- Snell-Hornby, M. 2006. *The turns of translation studies: new paradigms or shifting viewpoints?* Amsterdam: John Benjamins Publishing Company.
- Snover, M., Dorr, B., Schwartz, R., Micciulla, L., Makhoul, J. 2006. A study of translation edit rate with targeted human annotation. In: *AMTA 2006: Proceedings of the 7th Conference of the Association for Machine Translation in the Americas, "Visions for the Future of Machine Translation"*. Cambridge, Massachusetts, USA. 223-231.
- Somers, H. 2000. Machine Translation. In: Dale, R., Moisl, H., Somers, H. (eds.) 2000. *Handbook of Natural Language Processing*. Basel: Marcel Dekker. 329-346.
- Somers, H. L. 1998. Machine Translation, applications. In: Baker, M. (ed.) 1998. *Encyclopedia of Translation Studies*. London: Routledge. 136-140.
- Spearman, C. 1904. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*. Vol. 15. No. 1. Illinois: The University of Illinois Press. 72-101.

- Tellinger, D. 2007. A mértékegységek szerepe a műfordításban. *Alkalmazott Nyelvészeti Közlemények*. 2. évfolyam, 1. szám. Miskolc. 29-33.
- Thouin, B. 1981. The METEO system. In: *Practical experience of machine translation. Proceedings of a conference*. 1982. Lawson, V. (ed.) Amsterdam, New York, Oxford: North-Holland Publishing Company. 39-44.
- Tóth Beatrix 2006. A szövegértés fejlesztésének elmélete és gyakorlata. In: *Magyar Nyelvőr*. 130/4. 457–469.
- Tóth L. 2002. *Az olvasás pszichológiai alapjai*. Debrecen: Pedellus Tankönyvkiadó.
- Turian, J. P., Shen, L., Melamed, I. D. 2003. Evaluation of machine translation and its evaluation. In: *Proceedings of the MT Summit IX*. New Orleans, USA. 386-393.
- Vauquois, B. 1976. Automatic translation – a survey of different approaches. *Statistical Methods in Linguistics*. 1976. 127-135.

## 11. Mellékletek

### 11.1. A dolgozatban használt rövidítések

GF	gépi fordítás, MetaMorpho fordítása
GF_001	gépi fordítás, 1-es számú szöveg
EF	emberi fordítás
ESZ	eredeti, forrásnyelvi szöveg
ESZ_001	eredeti, 1-es számú szöveg
USZ	utószerkesztett szöveg
USZ_001	1-es számú utószerkesztett szöveg
MM	MetaMorpho fordítás ( <a href="http://www.webforditas.hu">www.webforditas.hu</a> )
GT	Google Translate fordítás
R	referenciális
RJ	referenciális jelentés
N	nyelvi
L	lexikai
S	stilisztikai
?	nem meghatározható szófaj
angol	FNY-ű szó
fn	főnév
fn, tul	tulajdonnév
fn/mn	főnévből –i képzővel képzett melléknév
hat	határozószó
hónap	hónap neve
ige	ige
ik	igekötő
in fn	főnévi igenév
in hat	határozói igenév
in mn	melléknévi igenév
ind	indulatszó
ír	írásjel
köt	kötőszó
mért	mértékegység
mn	melléknév
mód	módosítószó
nm	névmás
nm, ált	általános névmás
nm, birt	birtokos névmás
nm, hatlan	határozatlan névmás
nm, kérd	kérdő névmás

nm, kölcs	kölcsönös névmás
nm, mut	mutató névmás
nm, szem	személyes névmás
nm, tag	tagadó névmás
nm, utal	névmás utalószerkezetben
nm, vissza	visszaható névmás
nm, von	vonatkozó névmás
nu	névutó
nu - mn	melléknévi névutó
nvh	határozott névelő
nvl	határozatlan névelő
rag	rag
röv	rövidítés
szn	számnév
szn, év	évszám
szn, hat	határozott számnév
szn, hatlan	határozatlan számnév

## 11.2. A géppel fordított szövegek érthetőségét mérő kérdőív szövegei és kérdései

### 1. szöveg

#### BECKEK DRÓTVITÁBAN

DAVID Beckham bajban van a két méteres magas Colditze fölött-stílus akadály Beckingham Palace-nél.

A tanácsi tisztviselők azt mondják, hogy a Real Madrid csillagnak szüksége van arra, hogy engedély a szögesdrótnak felülmúlja a kerítést.

Ennek betolakodókat az udvarháztól távol kell tartania Sawbridgeworthnél, Herts, de a helyi lakosok szinkronizálták azt egy szemet sértő látvány.

Mondta egy szóvivő East Herts tanácsért: A kerítés rendben van de a drót szabályozásokat tör.

Egy tervezési alkalmazást csináltak de az információ eltűnt volt. Ők \*reapply(resupply).

A tanács megparancsolhatná neki, hogy vegye le a drótot.

#### Kérdések:

Miért került bajba David Beckham?

Mi a baj a drótkerítéssel?

### 2. szöveg

2006. március 17-e

#### A MAFFIA TÖBB EZRET RAGAD MEG

#### EXKLUZÍV INTERJÚ

*Stephen Moyes által*

Egy LEADING bukmékert tegnap akadályoztak rablók.

Freddie Williams, 63, ki dicsekedik azzal, hogy találkozóknál kezel több milliót, a lányával elhagyni Cheltenhamöt benne 20 volt, és egy másik ember volt.

Három ember balaclavas leállította a Jaguarját Cirencester közelében, és emelőrudakkal törte össze a Jaguarjuk ablakait.

Fenyegették Mr Williamsöt, Cumnock, Ayrshire, készpénzt adott át - mondott rendőrség által tíznek éves lenni több ezer.

A lánya rendőrséget hívott a színhelyből.

Az hármat vágott az, hogy üveget röptett, és mondtak lenni \*"traumatised"().

A banda két menekülő autóját tűzön találták.

#### Kérdések

Mi történt Freddie Williamsszel és a lányával?

Mit vettek el tőlük? Mennyit?

Mi lett velük?

### 3. szöveg

#### Nem tévedtem el. Csupán nagy felfedezőket emulálok

JANE SHILLING

A RAC biztosítótársasága éppen *Are We Lost Yet* kísértő címével kiadott egy felmérést?, melyik az, hogy a fiam mit mond nekem valami pontnál minden útnan, amit együtt vállalunk, pontosan elég furcsamód — általában azelőtt nekem sikerült az utunk a tetőjén megtárgyalnom a csomópontot.

A RAC felmérése egy arról a kedves öreg **szállítási gesztenyéről** végzett tanulmány, az emberek és nők különböző viselkedése, amikor **ez odajön ahhoz, hogy elveszék**. Tudod azt: a nők a természetfölöttien fejlett **népkeszségeikkel** ugranak a **hajtó ülésből** a pillanat meglehetősen Acacia Avenue-nál Acacia Grove-ban találják magukat és **hívás valami elhaladó helybelin eligazítani őket**. Az emberek, akik másfelől **sodródva megtalálják magukat**, megtagadják, hogy elismerjék, hogy ők bármilyen ilyen dolgok és dühös körökben száguldanak körbe miközben túráztatják a motort és összetörik a felszereléseket, a füleik going a **bíbor egy csúnya árnyéka** és a nyelvük, ami megdöbbenően leginkább megromlik. És nem világosan elhivatottak a belső sávba az m 1 észak felé tartó inkább mint az a 23 Brightonnek (melyik az, hogy róluk hol hiszik azt, hogy a kedvesükkel mennek egy nyugtató kiránduláson) el fogják ismerni, hogy van egy probléma. **Azáltal, hogy a kedves melyik időt üli meg az utasülésben, az ajkak annyira szorongtak egy vékony sorozatba nem adni kifejezést annak, hogy a szörnyű vélemények, amiket tartalmaztak a nagy gondolkodásléggömbben, lebegni a feje fölött.**

#### Kérdések

Miről szól az RAC felmérése?

Mit tesznek a nők, ha eltévednek?

Mit tesznek/milyenek a férfiak, ha eltévednek?

### 11.3. Az abszolút minimálisan utószerkesztett szövegek

#### 1. szöveg

2006. március 17-e

##### BECKEK DRÓTVITÁBAN

DAVID Beckham bajban van a két méteres magas szögesdrót kerítés miatt Beckingham Palace-nél.

A tanácsi tisztviselők azt mondják, hogy a Real Madrid csillagnak szüksége van arra, hogy engedély a szögesdrót legyen a kerítés fölött.

Ennek betolakodókat az udvarháztól távol kell tartania Sawbridgeworthnél, Herts, de a helyi lakosok mondták azt egy szemet sértő látványnak.

Mondta egy szóvivőnő East Herts tanácsért: A kerítés rendben van de a drót szabályozásokat tör.

Egy tervezési kérelmet csináltak de az információ hiányzott. Újra kell kérelmezniük.

A tanács megparancsolhatná neki, hogy vegye le a drótot.

#### 2. szöveg

2006. március 17-e

##### A MAFFIA TÖBB EZRET RAGAD MEG

##### EXKLUZÍV INTERJÚ

*Stephen Moyes által*

Egy menő bukmékert tegnap akadályoztak rablók.

Freddie Williams, 63, ki dicsekedik azzal, hogy találkozónál kezel több milliót, a lányával (20-as éveiben) és egy másik emberrel elhagyni Cheltenhamöt.

Három ember maszkban leállította a Jaguarját Cirencester közelében, és emelőrudakkal törte össze a Jaguarjuk ablakait.

Fenyegették Mr Williamsöt, Cumnockból, Ayrshire, aki készpénzt adott át - mondott rendőrség által lenni több tízezer.

A lánya rendőrséget hívott a színhelyből.

Az hármat vágott az, hogy üveg szétrepült, és mondtak lenni sokkos állapotban.

A banda két menekülő autóját kigyulladva találták.

##### Nem tévedtem el. Csupán nagy felfedezőket utánozok

JANE SHILLING

A RAC biztosítótársasága éppen az Are We Lost Yet? kísérteties címmel kiadott egy felmérést, amelyik pontosan elég furcsamód ugyanaz, amit a fiam mond nekem egy pontnál minden útnál, amit együtt vállalunk, — általában mielőtt nekem sikerült az utunk a tetőjén megértenem a csomópontot.

A RAC felmérése egy arról a kedves öreg közlekedési történetről végzett tanulmány, a férfiak és nők különböző viselkedése, amikor odajutnak, hogy elvesszék. Tudod azt: a nők a természetfölöttien fejlett kommunikációs készségeikkel azonnal ugranak a hajtó ülésből amikor Acacia Avenue helyett Acacia Grove-ban találják magukat és hívás valami elhaladó helybelin eligazítani őket. Az emberek, akik másfelől eltévedve találják magukat, megtagadják, hogy elismerjék, hogy ők bármilyen ilyen dolgok és dühös körökben száguldanak körbe miközben túráztatják a motort és összetörik a felszereléseket, a füleik a bíbor egy csúnya árnyéka lesz és a nyelvük, ami megdöbbenően leginkább megromlik. És amíg nem világosan tudják, hogy a belső sávba az M1 észak felé tartó irányban haladnak inkább mint az A23 Brighton felé (melyikről hiszik azt, hogy a kedvesükkel oda mennek egy nyugtató kiránduláson) nem fogják elismerni, hogy van egy probléma. Addigra a kedves melyik időt üli meg az utasülésben, az ajkak annyira szorongtak egy vékony sorozatba nem adni kifejezést annak, hogy a szörnyű vélemények, amiket tartalmaztak a nagy gondolkodásleggömbben, lebegni a feje fölött.



## 11.4. Korpusz

A vizsgálatokhoz felhasznált korpusz teljes anyaga megtalálható a CD mellékleten a Korpusz mappa Eredeti\_szovegek, Gepi\_forditas és Utszerkesztett nevű almappáiban. A mappákban található fájlok neve a dolgozatban is hivatkozásként használt számokat tartalmazza. A példamondatokban a mondatok után rövidítéssel és a szöveg számával hivatkozunk a mondatok forrására.

A fájlok nevének listája:

## 11.5. 10 szóból álló szegmens összes lehetséges felbontása

Meteor sorrend	Új sorrend									Fedett- ség	ME- TEOR arány	METEOR büntetés	Végső képlet	Büntetés
1.	1.	10								10	0,100	0,000019	0,0250	0,000008
2.	14.	9								9	0,111	0,000182	0,0789	0,000246
3.	44.	8								8	0,125	0,000768	0,1389	0,001340
4.	81.	7								7	0,143	0,002315	0,2059	0,004363
5.	103.	6								6	0,167	0,005870	0,2813	0,011124
6.	118.	5								5	0,200	0,013500	0,3667	0,024648
7.	5.	4	6							10	0,200	0,000148	0,0435	0,000041
8.	6.	5	5							10	0,200	0,000148	0,0444	0,000044
9.	2.	1	9							10	0,200	0,000148	0,0408	0,000034
10.	3.	2	8							10	0,200	0,000148	0,0417	0,000036
11.	4.	3	7							10	0,200	0,000148	0,0426	0,000039
12.	16.	1	8							9	0,222	0,000615	0,0870	0,000329
13.	20.	3	6							9	0,222	0,000615	0,0909	0,000376
14.	24.	4	5							9	0,222	0,000615	0,0930	0,000402
15.	126.	4								4	0,250	0,029407	0,4643	0,050041
16.	45.	1	7							8	0,250	0,001821	0,1395	0,001358
17.	49.	2	6							8	0,250	0,001821	0,1429	0,001458
18.	53.	3	5							8	0,250	0,001821	0,1463	0,001567
19.	56.	4	4							8	0,250	0,001821	0,1500	0,001688
20.	75.	1	6							7	0,286	0,004521	0,2000	0,004000
21.	78.	2	5							7	0,286	0,004521	0,2051	0,004316
22.	82.	3	4							7	0,286	0,004521	0,2105	0,004665
23.	7.	1	1	8						10	0,300	0,000554	0,0625	0,000122
24.	8.	1	3	6						10	0,300	0,000554	0,0652	0,000139
25.	9.	2	2	6						10	0,300	0,000615	0,0652	0,000139
26.	10.	1	4	5						10	0,300	0,000554	0,0667	0,000148
27.	11.	2	3	5						10	0,300	0,000615	0,0667	0,000148
28.	12.	2	4	4						10	0,300	0,000615	0,0682	0,000158
29.	13.	3	3	4						10	0,300	0,000686	0,0682	0,000158
30.	132.	3								3	0,333	0,062500	0,5769	0,096012
31.	99.	1	5							6	0,333	0,010143	0,2703	0,009871
32.	101.	2	4							6	0,333	0,010143	0,2778	0,010177
33.	105.	3	3							6	0,333	0,010143	0,2857	0,011662
34.	28.	1	1	7						9	0,333	0,001626	0,1111	0,000686
35.	30.	1	2	6						9	0,333	0,001626	0,1136	0,000734
36.	33.	1	3	5						9	0,333	0,001626	0,1163	0,000786
37.	34.	2	2	5						9	0,333	0,001821	0,1163	0,000786
38.	37.	1	4	4						9	0,333	0,001626	0,1190	0,000844
39.	38.	2	3	4						9	0,333	0,001821	0,1190	0,000844
40.	40.	3	3	3						9	0,333	0,002048	0,1220	0,000907
41.	60.	1	1	6						8	0,375	0,004000	0,1667	0,002315
42.	63.	1	2	5						8	0,375	0,004000	0,1707	0,002488
43.	66.	1	3	4						8	0,375	0,004000	0,1750	0,002680

Meteor sorrend	Új sorrend									Fedett- ség	ME- TEOR arány	METEOR büntetés	Végső képlet	Büntetés
44.	67.	2	2	4						8	0,375	0,004521	0,1750	0,002680
45.	69.	2	3	3						8	0,375	0,004521	0,1795	0,002891
46.	115.	1	4							5	0,400	0,021438	0,3529	0,021982
47.	117.	2	3							5	0,400	0,021438	0,3636	0,024042
48.	15.	1	1	1	7					10	0,400	0,001458	0,0851	0,000308
49.	17.	1	1	2	6					10	0,400	0,001458	0,0870	0,000329
50.	18.	1	1	3	5					10	0,400	0,001458	0,0889	0,000351
51.	19.	1	2	2	5					10	0,400	0,001626	0,0889	0,000351
52.	21.	1	1	4	4					10	0,400	0,001458	0,0909	0,000376
53.	22.	1	2	3	4					10	0,400	0,001626	0,0909	0,000376
54.	23.	2	2	2	4					10	0,400	0,001821	0,0909	0,000376
55.	25.	1	3	3	3					10	0,400	0,001821	0,0930	0,000402
56.	26.	2	2	3	3					10	0,400	0,001821	0,0930	0,000402
57.	88.	1	1	5						7	0,429	0,008876	0,2308	0,006145
58.	90.	1	2	4						7	0,429	0,008876	0,2368	0,006643
59.	92.	1	3	3						7	0,429	0,008876	0,2432	0,007196
60.	93.	2	2	3						7	0,429	0,010143	0,2432	0,007196
61.	42.	1	1	1	6					9	0,444	0,003556	0,1364	0,001268
62.	46.	1	1	2	5					9	0,444	0,003556	0,1395	0,001358
63.	50.	1	1	3	4					9	0,444	0,003556	0,1429	0,001458
64.	51.	1	2	2	4					9	0,444	0,004000	0,1429	0,001458
65.	54.	1	2	3	3					9	0,444	0,004000	0,1463	0,001567
66.	55.	2	2	2	3					9	0,444	0,004521	0,1463	0,001567
67.	134.	2								2	0,500	0,132835	0,7083	0,177698
68.	125.	1	3							4	0,500	0,043896	0,4516	0,046054
69.	127.	2	2							4	0,500	0,043896	0,4667	0,050815
70.	108.	1	1	4						6	0,500	0,018519	0,3056	0,014264
71.	110.	1	2	3						6	0,500	0,018519	0,3143	0,015522
72.	111.	2	2	2						6	0,500	0,021438	0,3235	0,016932
73.	73.	1	1	1	5					8	0,500	0,007813	0,1951	0,003714
74.	76.	1	1	2	4					8	0,500	0,007813	0,2000	0,004000
75.	79.	1	1	3	3					8	0,500	0,007813	0,2051	0,004316
76.	80.	1	2	2	3					8	0,500	0,008876	0,2051	0,004316
77.	83.	2	2	2	2					8	0,500	0,010143	0,2105	0,004665
78.	27.	1	1	1	1	6				10	0,500	0,003175	0,1087	0,000642
79.	29.	1	1	1	2	5				10	0,500	0,003175	0,1111	0,000686
80.	31.	1	1	1	3	4				10	0,500	0,003175	0,1136	0,000734
81.	32.	1	1	2	2	4				10	0,500	0,003556	0,1136	0,000734
82.	35.	1	1	2	3	3				10	0,500	0,003556	0,1163	0,000786
83.	36.	1	2	2	2	3				10	0,500	0,004000	0,1163	0,000786
84.	39.	2	2	2	2	2				10	0,500	0,004521	0,1190	0,000844
85.	58.	1	1	1	1	5				9	0,556	0,006912	0,1628	0,002157
86.	61.	1	1	1	2	4				9	0,556	0,006912	0,1667	0,002315
87.	64.	1	1	1	3	3				9	0,556	0,006912	0,1707	0,002488
88.	65.	1	1	2	2	3				9	0,556	0,007813	0,1707	0,002488
89.	68.	1	2	2	2	2				9	0,556	0,008876	0,1750	0,002680

Meteor sorrend	Új sorrend										Fedett- ség	ME- TEOR arány	METEOR büntetés	Végső képlet	Büntetés
90.	97.	1	1	1	4						7	0,571	0,016106	0,2632	0,009112
91.	100.	1	1	2	3						7	0,571	0,016106	0,2703	0,009871
92.	102.	1	2	2	2						7	0,571	0,018519	0,2778	0,010717
93.	121.	1	1	3							5	0,600	0,037323	0,3939	0,030567
94.	122.	1	2	2							5	0,600	0,037323	0,4063	0,033524
95.	41.	1	1	1	1	1	5				10	0,600	0,006145	0,1333	0,001185
96.	43.	1	1	1	1	2	4				10	0,600	0,006145	0,1364	0,001268
97.	47.	1	1	1	1	3	3				10	0,600	0,006145	0,1395	0,001358
98.	48.	1	1	1	2	2	3				10	0,600	0,006912	0,1395	0,001358
99.	52.	1	1	2	2	2	2				10	0,600	0,007813	0,1429	0,001458
100.	86.	1	1	1	1	4					8	0,625	0,014096	0,2250	0,005695
101.	89.	1	1	1	2	3					8	0,625	0,014096	0,2308	0,006145
102.	91.	1	1	2	2	2					8	0,625	0,016106	0,2368	0,006643
103.	131.	1	2								3	0,667	0,088989	0,5714	0,093294
104.	114.	1	1	1	3						6	0,667	0,032000	0,3429	0,020152
105.	116.	1	1	2	2						6	0,667	0,032000	0,3529	0,021982
106.	71.	1	1	1	1	1	4				9	0,667	0,012406	0,1905	0,003455
107.	74.	1	1	1	1	2	3				9	0,667	0,012406	0,1951	0,003714
108.	77.	1	1	1	2	2	2				9	0,667	0,014096	0,2000	0,004000
109.	57.	1	1	1	1	1	1	4			10	0,700	0,010976	0,1591	0,002013
110.	59.	1	1	1	1	1	2	3			10	0,700	0,010976	0,1628	0,002157
111.	62.	1	1	1	1	2	2	2			10	0,700	0,012406	0,1667	0,002315
112.	107.	1	1	1	1	3					7	0,714	0,027643	0,2973	0,013138
113.	109.	1	1	1	2	2					7	0,714	0,027643	0,3056	0,014264
114.	129.	1	1	2							4	0,750	0,074191	0,5000	0,062500
115.	96.	1	1	1	1	1	3				8	0,750	0,024042	0,2564	0,008429
116.	98.	1	1	1	1	2	2				8	0,750	0,024042	0,2632	0,009112
117.	85.	1	1	1	1	1	1	3			9	0,778	0,021041	0,2195	0,005289
118.	87.	1	1	1	1	1	2	2			9	0,778	0,021041	0,2250	0,005695
119.	124.	1	1	1	2						5	0,800	0,062500	0,4375	0,041870
120.	70.	1	1	1	1	1	1	1	3		10	0,800	0,018519	0,1860	0,003220
121.	72.	1	1	1	1	1	1	2	2		10	0,800	0,018519	0,1905	0,003455
122.	120.	1	1	1	1	2					6	0,833	0,053142	0,3824	0,027949
123.	113.	1	1	1	1	1	2				7	0,857	0,045563	0,3333	0,018519
124.	106.	1	1	1	1	1	1	2			8	0,875	0,039359	0,2895	0,012128
125.	95.	1	1	1	1	1	1	1	2		9	0,889	0,034232	0,2500	0,007813
126.	84.	1	1	1	1	1	1	1	1	2	10	0,900	0,029958	0,2143	0,004920
127.	136.	1									1	1,000	0,289352	0,8636	0,322079
128.	135.	1	1								2	1,000	0,182216	0,7200	0,186624
129.	133.	1	1	1							3	1,000	0,148148	0,6296	0,124803
130.	130.	1	1	1	1						4	1,000	0,122070	0,5517	0,083972
131.	128.	1	1	1	1	1					5	1,000	0,101771	0,4839	0,056645
132.	123.	1	1	1	1	1	1				6	1,000	0,085734	0,4242	0,038178
133.	119.	1	1	1	1	1	1	1			7	1,000	0,072897	0,3714	0,025621
134.	112.	1	1	1	1	1	1	1	1		8	1,000	0,062500	0,3243	0,017057
135.	104.	1	1	1	1	1	1	1	1	1	9	1,000	0,053990	0,2821	0,011219
136.	94.	1	1	1	1	1	1	1	1	1	10	1,000	0,046957	0,2439	0,007255

## 11.6. Az optimális felbontást megtaláló algoritmus C nyelvű kódhoz közeli pszeudokódja

Az intervallumban található szakaszoknak csak a kezdőpontját és hosszát tároljuk egy tömbben: a tömb indexe a szakasz kezdőpontja, a tömb eleme pedig a szakasz hossza. A felbontást biztosító szakaszoknak csak a hossza érdekes, kezdőpontjukat nem tároljuk. A sorrendjük sem fontos, de az algoritmus a sorrendet megtartja. A kezdőpontok tárolásához az algoritmus könnyen módosítható. A felbontást tartalmazó szakaszokat szintén tömbben tároljuk.

MAXHOSSZ - a szakaszok maximális hossza  
MAXSZOSZAM - az intervallum (szegmens) maximális hossza  
mondat[MAXSZOSZAM] -- egész számok, az intervallumban található szakaszok hosszai  
MondatFelbontas[MAXSZOSZAM] -- egész számok, a fedést biztosító szakaszok hosszai, azaz a felbontás  
SzakaszSzam -- a felbontásban szereplő szakaszok száma

FelbontasHatra függvény: utolsoindextől kezdve visszafelé haladva elsoindexig a szakaszok[] tömbbe teszi a felbontáshoz szükséges szakaszokat a megfelelő sorrendben.  
paraméterek:  
mondat[] - egész számok, az intervallum szakaszai  
elsoindex - egész szám, az intervallum kezdőpontja  
utolsoindex - egész szám, az intervallum végpontja  
szakaszok[] - egész számok, a felbontás szakaszai  
A függvény visszatérési értéke: a felbontásban szereplő szakaszok száma.

Lokális változók:  
k - egész szám, a mondat[] végigjárásához használt ciklusváltozó  
s - egész szám, a szakaszokon belüli vizsgálathoz használt ciklusváltozó  
x - egész szám, a szakaszok[] tömb indexeléséhez használt ciklusváltozó  
leghosszabbindex - egész szám, az aktuális elemet még lefedő leghosszabb szakasz pozíciója

Megjegyzés: Minden tömb indexelése 1-től kezdődik. Tömbök esetében cím szerinti paraméterátadást feltételezünk.

```
Function FelbontasHatra(mondat[], elsoindex, utolsoindex, szakaszok[])
{
    k = utolsoindex /* a vizsgálatot az intervallum vegerol kezdjuk */
    x = 1
    while (k >= elsoindex) /* az elejeig megyunk a mondatnak */
    {
        leghosszabbindex = 0
        /* ha nem találunk olyan szakaszt, amely lefedi a k-adik
        elemet, akkor 0 marad */
        s = k
        /* megkeressük az aktuális elemet lefedő leghosszabb szakaszt
        */
        while ((s >= (k - MAXHOSSZ + 1)) && (s >= 1))
        {
```

```

        if(s + mondat[s] > k && mondat[s] > 0 && s >= elsoindex)
            leghosszabbindex = s
        s = s - 1
    }
    /* ha volt olyan szakasz, amely lefedi az aktuális elemet */
    if(leghosszabbindex > 0)
    {
        szakaszok[x] = k - leghosszabbindex + 1
        x = x + 1
        k = k - szakaszok[x - 1] /* a következő vizsgálandó
        elemre ugrunk */
    }
    else
        k = k - 1
    }
    /* az eddig fordított sorrendben megkapott, a felbontást tartalmazó
    tömb elemeinek sorrendjének megfordítása */
    /* paraméterei: a megfordítandó tömb és az elemszám */
    MegforditTomb (szakaszok, x - 1)
    return x - 1
}

```

A FelbontasKeszit függvény elkészíti a legjobb felbontást, amellyel a MondatFelbontas tömböt tölti fel.  
 A függvény visszatérési értéke a felbontásban szereplő szakaszok száma.  
 A paraméterek:  
 mondat[] - egész számok, az intervallum szakaszai  
 MondatFelbontas[] - egész számok, a felbontás szakaszai  
 n - egész szám, az intervallum hossza

Lokális változók:

Minden változó típusa egész szám, ha nincs külön jelölve.  
 k - a mondat[] végigjárásához használt ciklusváltozó  
 s - a szakaszokon belüli vizsgálathoz használt ciklusváltozó  
 x - a MondatFelbontas[] tömb indexeléséhez használt ciklusváltozó  
 hossz - az aktuális elem hossza, nem feltétlenül szükséges, hiszen megegyezik a mondat[k] értékkel  
 kovhossz - a szakasz után azonnal következő szakasz hossza  
 maxtindex - a szakaszon belüli, a szakasz végpontját legnagyobb mértékben túllépő szakasz indexe  
 maxthossz - a túllépés mértéke  
 vanhosszabb - logikai változó, azt jelzi, hogy van-e egyáltalán ilyen hosszabb szakasz  
 akteloreindex - a belső, az aktuális szakasznál hosszabb, de nem a kellő mértékben túllépő szakasz indexe !!!ezeknek a nevét meg kell változtatni!!!  
 akteloreelem - az elem hossza !!kell egyáltalán?  
 vambenneagyobb - logikai változó, addig igaz, amíg van ilyen, az aktuális szakasznál hosszabb szakasz, minden szakaszra  
 nincstobbnagyobb - logikai változó, azt jelzi, hogy az aktuálisan vizsgált szakaszban van-e még saját magánál hosszabb szakasz  
 maxbenneindex -  
 hanyhatraszakasz - a visszafelé haladó algoritmus hány darab szakaszt talált, amelyet fel kell vennünk a fedésbe  
 honnanvissza - a visszafelé haladó algoritmus kezdőpontja, amelyet a FelbontasHatra függvénynek paraméterként átadunk  
 eddighatraindex - a visszafelé haladó algoritmus végpontja, amelyet a FelbontasHatra függvénynek paraméterként átadunk  
 hatraszakaszok[] - a FelbontasHatra függvénynek paraméterként átadott tömb, amelybe a felbontásban szereplő szakaszokat teszi  
 i - ciklusváltozó

```

Function FelbontasKeszit(mondat[], MondatFelbontas[], n)
{
    maxthossz = 0
    vanhosszabb = False
    nincstobbnagyobb = True
    vanbennenagyobb = False
    k = 1
    x = 1
    while( k <= n )
    {
        if( mondat[k] = 0 )
            /* ha az aktuális elemtől nem kezdődik szakasz */
            {
                k = k + 1
            }
        else if( vanhosszabb == False && mondat[k] == 1 )
        {
            /* az elejétől, vagy ha teljesen új szakaszt kezdünk (nem
            belső szakaszt), akkor az 1 hosszúságú szakaszokat
            felvesszük a felbontásba */
            MondatFelbontas[x] = mondat[k]
            x = x + 1
            k = k + 1
        }
        else if( k + mondat[k] == n )
        {
            /* ha vége van a mondatnak */
            mondatfelbontas[x] = mondat[k]
            k = k + mondat[k]
            x = x + 1
        }
        else if( (k + mondat[k] > n) )
        {
            /* az utolsó szakasz túllépi az intervallum végpontját -
            gyakorlatban nem fordulhat elő */
            MondatFelbontas[x] = n - k + 1
            k = k + mondat[k]
            x = x + 1
        }
        else
        {
            if( vanhosszabb == False )
            {
                /* előlről kezdtük a szakaszt, amelynek hossza
                nagyobb, mint 1 */
                hossz = mondat[k]
            }
            else
            {
                /* volt az aktuális szakaszban túllépő szakasz */
                hossz = mondatfelbontas[x]
            }
            s = 1 /* a szakaszon belüli indexeles */
            kovhossz = mondat[k + hossz]
            maxthossz = 0
            maxtindex = 0
            vanhosszabb = False
        }
    }
}

```

```

while( s < hossz)
{
    /* van-e a végpontot túllépő szakasz */
    if( mondat(k + s) > hossz - s + kovhossz)
    {
        vanhosszabb = True
        if( maxthossz < mondat[k+s] - (hossz-s))
        {
            maxthossz = mondat[k+s] - (hossz-s)
            maxtindex = k + s
        }
    }
    s = s + 1
}

/* szakaszok tárolása */
if( vanhosszabb == True && hossz >= mondat[maxtindex])
{
    mondatfelbontas[x] = hossz
    x = x + 1
    mondatfelbontas[x] = maxthossz
    k = k + MondatFelbontas[x - 1]
}
else if( vanhosszabb == True And hossz < maxthossz)
{
    mondatfelbontas[x] = maxtindex - k
    x = x + 1
    mondatfelbontas[x] = mondat[maxtindex]
    k = k + MondatFelbontas[x - 1]
}
else if( vanhosszabb == False)
{
    /* ha nem volt megfelelően túllépő szakasz, meg kell
    vizsgálni, hogy van-e az aktuális szakasznál hosszabb belső
    szakasz */
    eddighatraindex = k
    akteloreindex = k
    /* a hatraszakaszok[] tömböt le kell nullázni */
    for( i = 1 ; i <= MAXSZOSZAM; i++)
        hatraszakaszok[i] = 0
    vanbinnenagyobb = False
    /* ebben a szakaszszorozatban van-e hosszabb belső
    szakasz */
    nincstobbnagyobb = False
    /* az aktuális szakaszban van-e hosszabb belső szakasz
    */
    maxbenneindex = 0
}

```



```

while(akteloreindex <= n && nincstobbnagyobb == False)
{
    /* addig vizsgáljuk a szakaszokat, amíg van bennük
    hosszabb szakasz */
    /* a végén az utolsó ilyen szakasz kezdete:
    akteloreindex */
    for( t = akteloreindex + 1 ; t<=akteloreindex +
    mondat[akteloreindex] - 1 ; t++)
    {
        /* az aktuális szakaszon belüli leghosszabb
        szakasz kiválasztása */
        if( mondat[i] >= mondat[maxbenneindex]
        {
            maxbenneindex = i
        }
    }
    if( mondat[maxbenneindex]>mondat[akteloreindex])
    {
        /* ha az előbb kiválasztott leghosszabb szakasz
        hosszabb, mint az aktuális szakasz */
        vanbennenagyobb = True
        akteloreindex = maxbenneindex
    }
    else
    {
        /* ebben az aktuális szakaszban nincs magánál
        hosszabb szakasz */
        nincstobbnagyobb = True
    }
}
if( vanbennenagyobb == True )
{
    honnanvissza = maxbenneindex +
    mondat[maxbenneindex] - 1
    /* hátrafelé haladó vizsgálat megkezdése */
    hanyhatraszakasz = FelbontasHatra(mondat,
    eddighatraindex, honnanvissza, hatraszakaszok)
    i = 1
    while( i <= hanyhatraszakasz)
    /* a hátrafelé haladó algoritmusból származó
    szakaszok bemásolása a felbontások közé */
    {
        MondatFelbontas[x] = hatraszakaszok[i]
        x = x + 1
        i = i + 1
    }
    k = honnanvissza + 1
}
else
/* ha nem volt túllépő szakasz vagy belső, hosszabb
szakasz */
{
    MondatFelbontas[x] = hossz
    x = x + 1
    k = k + MondatFelbontas[x - 1]
}
}
return x - 1
}

```